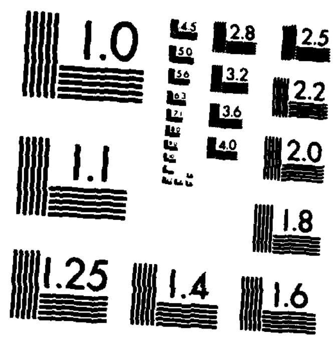
AD-A145 561 EFFECTS OF ALTERNATIVE CHROMATIC MIXED GRAPHICS DISPLAYS IN DECISION SUPPORT SYSTEMS(U) AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH M S MCCULLY MAY 84 1/4

UNCLASSIFIED AFIT/CI/NR-84-46D F/G 9/2 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

UNCLASS

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

AD-A145 561

DTIC FILE COPY

(1)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AFIT/CI/NR 84-46D	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): Effects of Alternative Chromatic Mixed Graphics Displays in Decision Support Systems	5. TYPE OF REPORT & PERIOD COVERED: ~~THESIS~~/DISSERTATION
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Mary Susan McCully	8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS: AFIT STUDENT AT: Arizona State University	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS: AFIT/NR WPAFB OH 45433	12. REPORT DATE: May 1984
	13. NUMBER OF PAGES: 270
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): UNCLASS
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

DTIC ELECTE SEP 17 1984 S D B

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

APPROVED FOR PUBLIC RELEASE: IAW AFR 190-17

24 July 84

Lynn Wolaver
LYNN E. WOLAVER
Dean for Research and Professional Development
AFIT, Wright-Patterson AFB OH

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

ATTACHED

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASS

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

84 09 13 011

AFIT RESEARCH ASSESSMENT

The purpose of this questionnaire is to ascertain the value and/or contribution of research accomplished by students or faculty of the Air Force Institute of Technology (AU). It would be greatly appreciated if you would complete the following questionnaire and return it to:

AFIT/NR
Wright-Patterson AFB OH 45433

RESEARCH TITLE: Effects of Alternative Chromatic Mixed Graphics Displays in Decision Support Systems

AUTHOR: Mary Susan McCully

RESEARCH ASSESSMENT QUESTIONS:

1. Did this research contribute to a current Air Force project?

( ) a. YES ( ) b. NO

2. Do you believe this research topic is significant enough that it would have been researched (or contracted) by your organization or another agency if AFIT had not?

( ) a. YES ( ) b. NO

3. The benefits of AFIT research can often be expressed by the equivalent value that your agency achieved/received by virtue of AFIT performing the research. Can you estimate what this research would have cost if it had been accomplished under contract or if it had been done in-house in terms of manpower and/or dollars?

( ) a. MAN-YEARS ________ ( ) b. $________

4. Often it is not possible to attach equivalent dollar values to research, although the results of the research may, in fact, be important. Whether or not you were able to establish an equivalent value for this research (3. above), what is your estimate of its significance?

( ) a. HIGHLY SIGNIFICANT ( ) b. SIGNIFICANT ( ) c. SLIGHTLY SIGNIFICANT ( ) d. OF NO SIGNIFICANCE

5. AFIT welcomes any further comments you may have on the above questions, or any additional details concerning the current application, future potential, or other value of this research. Please use the bottom part of this questionnaire for your statement(s).

NAME GRADE POSITION

ORGANIZATION LOCATION

STATEMENT(s):

FOLD DOWN ON OUTSIDE - SEAL WITH TAPE

AFIT/NR
WRIGHT-PATTERSON AFB OH 45433

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE. $300

NO POSTAGE
NECESSARY
IF MAILED
IN THE
UNITED STATES

BUSINESS REPLY MAIL
FIRST CLASS PERMIT NO. 73236 WASHINGTON D.C.

POSTAGE WILL BE PAID BY ADDRESSEE

AFIT/ DAA
Wright-Patterson AFB OH 45433

FOLD IN

EFFECTS OF ALTERNATIVE CHROMATIC MIXED GRAPHICS DISPLAYS

IN

DECISION SUPPORT SYSTEMS

by

Mary Susan McCully

A Dissertation Presented In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

ARIZONA STATE UNIVERSITY

May 1984

EFFECTS OF ALTERNATIVE CHROMATIC MIXED GRAPHICS DISPLAYS

IN

DECISION SUPPORT SYSTEMS

by

Mary Susan McCully

has been approved

March 1984

APPROVED:

_________________________, Chairperson

____

____

____

____

Supervisory Committee

ACCEPTED:

____
Department Chairperson

____
Dean, Graduate College

ABSTRACT

The relative effectiveness on decision makers' performance when integrating color functionally into the mixed graphics of a Decision Support System (DSS) was investigated. The method of research was a laboratory approach using a man-machine experiment. The key factors of the experimental design were determined by the components of Mason and Mitroff's (1973) definition of a DSS. The decision makers had academic engineering backgrounds. The performance measures included 3 levels of decision time and 12 levels of decision quality. Seven characteristics of decision makers were included: age, sex, college level and experience, attitude toward the computer, and experience with computers and video games. Five alternatives of colors were functionally integrated into mixed graphics (text and graphics) displays: four chromatic and one monochrome. The decisions represented a class of problems which were structured and had deterministic data. The interface employed was computer generated hardcopy. The organizational context was managerial control of resources.

One hundred eleven hypotheses were formulated. The data collected on the performance of 120 subjects who

worked an hour long engineering inventory case study were rigorously analyzed to gain insight on the effects of color.

Of the 86 models which measured decision quality, 78 contained insufficient evidence to conclude chromatic graphics displays differentially affected performance. The typically used CRT colors correlated with the best performance in the majority of the models which supported differential performance. This CRT alternative averaged a 9.2% nonsignificant increase in decision quality over the second place monochrome alternative. Only 1 of the 25 models which measured time contained sufficient evidence to conclude color treatment differentially affected performance. A chromatic alternative ranked first, supporting a nonsignificant 2.7% increase in performance over the second ranked monochrome alternative.

These results, though inconsistent with the theoretical and subjective reports, are concurrent with other related empirical research. Color does not have a unique quality for functionally coding information in a DSS. If color is incorporated into a DSS, those colors currently preset on many CRTs appear best. However increased efficiency in a decision maker's performance attributable to functional color can not be assumed.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

LIST OF TABLES

LIST OF FIGURES

I. THE PROBLEM AND ITS BACKGROUND

Introduction

Computer generated graphics supporting the option of color have become the heart of many decision support systems (DSS) in the industrial, business and military sectors. Technological advances in the microelectronic industry, an increasing reliance on computers by decision makers, and the affordability, availability and usability of the hardware, software and servicing have made these systems proliferate steadily. However these systems

> have been growing seemingly in an ad hoc manner. While support and praise of these systems has been relatively high, it appears that efforts have fallen short of realizing maximum effectiveness and efficiency for the system designed (Dickson, Senn and Chervany, 1977, p. 913).

The technical development of the color graphical displays is ahead of reported, quantitative information regarding its contribution to the DSS. The challenge remains to ascertain its relative effectiveness on user performance (Christ and Corso, 1983; Durrett, 1979; Haber and Wilkinson, 1982; Hanson, 1979; Leith and O'Shea, 1980; Riesenfeld, 1978; Tullis, 1981; Wallace, 1981).

Background

Color coding has been widely used as a method of information input code for visual displays for over 25

years. However, empirical evidence and guidelines on how to use it effectively are sparse in the literature with the reports very narrow in scope. In fact the literature has just begun to address how color impacts DSS.

After an extensive review and analysis of the literature on color in visual displays, the literature was categorized under one of three headings: theoretical support, subjective reasoning and objective experimental findings. The theoretical and subjective literature are concurrent in their reports: color allows for relatively more efficient processing of information, thus improves performance over use of a *monochrome* display (DeMars, 1975; Friend, 1980; Miller, 1982; Morris, 1979; Truckenbroad, 1981).

The majority of the empirical studies based their conclusions on analysis of performance of the simple discrete tasks of search, locate, identify and count type tasks. Memory was included in only three studies. The environment was an isolated laboratory setting that often simulated an aircraft cockpit. The alternative codes to which color was compared for relative effectiveness included letters, numerals, geometric and abstract shapes. The subjects, whose performance constituted the data for analysis, were usually lacking in experience with the

given task. Without exception the measure of performance was speed and/or accuracy.

Given these conditions, the findings in the experimental literature are not consistent with that from the theoretical and subjective reports. The objective results can be summarized by stating the relative effectiveness of color is a function of the task undertaken, the other display conditions and the measure of performance. Generally, color was not found to have a unique quality for coding information in a display over other codes (Barker and Krebs, 1977; Christ, 1977; Christner and Ray, 1961; Cook, 1974; Tullis, 1981).

Gaps in Knowledge

There continues to remain glaring gaps in the literature on the use of color in visual displays, especially those which do not involve search, count, locate and identify types of tasks. With the computer graphics penetrating almost every sector of the technological world, visual displays are being used everywhere for a variety of activities which involve information processing. One of the fastest growing sectors of users are managers and decision makers who are allocating their resources on incorporating color graphics into their DSS (Berk, Brownston and Kaufman, 1982; McLamb,

1982). Will color on their visual displays aid managers in their primary function of making decisions at different levels? The subjective literature purports it will. It advocates color will speed up the decision making process significantly as well as allow for better management recall or memory, thus better decisions (Durrett, Zwiener and Freund, 1981; Friend, 1980; Miller, 1982; Morris, 1979). For most decision makers time and accuracy are two of their most coveted resources (Kelley, 1979; McLamb, 1982; Miastkowski and Wrege, 1981; Takeuchi and Schmidt, 1980; Watson and Driver, 1983; Zmud, 1979).

However, this question of whether color does allow for quicker and/or better decisions has not been rigorously addressed, thus remains unanswered. Only two studies were found that involved color and graphics in a information system (IS) environment. Dooley and Harkins (1970) investigated the effect color had on attention and memory, when incorporated into a graphics display that sat in the corner of the experiment room. The subjects did not realize the graphics display was part of the experiment until it was removed from the room and the post test administered to collect memory data. The data from this study contained insufficient evidence to conclude color affected memory performance. However, it did affect

attention time.

Tullis (1981) had his subjects perform a training and diagnostic task involving the status of telephone lines. The graphics were line schematics displayed on a CRT interface. Conditions of the telephone lines were coded by either an achromatic shape or chromatic underlining bar (red, yellow, green). The task involved decision making on a low level and no memory. Based on his analysis of the data, he concluded that there was no significant differences in time or accuracy performance between the two graphics conditions of the color underlining bar and the achromatic shapes.

Problem Statement

Color graphics technology is here and is being used extensively in DSS. However there still remains much to be learned regarding when and how it can be used to increase a decision maker's performance. To realize what benefits are gained by incorporating color into the graphics displays of a DSS, quantitative information is needed on the relative effectiveness color has on the performance of the decision maker. The research reported here was designed in attempts to narrow this void of quantitative information on the relative impact chromatic graphics displays have on the performance of the decision

maker.

Specifically, the objective of the reported research was to gain insight on how the attribute of color, when integrated functionally into a DSS, affects the performance of a decision maker. The DSS under investigation involved decision makers with academic engineering backgrounds. The decisions represented a general class of problems involving managerial control of resources. The mode of presentation was mixed graphics which were displayed on computer generated hardcopy.

Research Approach and Dissertation Outline

This chapter concludes with the highlights of each of the chapters which follow. The research approach is developed throughout this outline of the dissertation.

Literature Search (Chapter II)

Initially the literature was searched to determine what had been reported in the area of the impact color displays had on human performance. This was accomplished without preference to type of performance measured, visual display interface used or chromatic and achromatic codes compared. Studies based on theoretical, subjective and objective analyses were all included.

The Management Information System (MIS) literature was also searched for two reasons. First, it was examined

to glean what had been reported on the topic of color and/or graphics displays and DSS. Second it was scrutinized to discover a taxonomy that had been validated and recommended for use in DSS research.

Based on the void discovered in the literature in the quantitative knowledge of the effects the technology of color has on the performance of a decision maker with an engineering background, the problem statement was formulated. Chapter II, Literature Search, delineates and summarizes the studies reviewed, defines the DSS and addresses the areas where more research is needed.

Methodology (Chapter III)

The MIS literature recommends research be developed through many small studies (Lucas, 1981) and in particular done in a laboratory environment which allows for controls (Dickson, Senn and Chervany, 1977). Since the objectives of this research could be met within a framework of a laboratory approach, an experiment was designed using this approach.

The key parts of a DSS given in Mason and Mitroff's (1973) definition are a person of a certain psychological type, a class of problems, data of a certain form, a particular interface and an organizational setting. The DSS investigated in this study was composed of decision

makers with academic engineering backgrounds. They solved a basic engineering case study which involved structured problems and deterministic data. The data were presented in mixed graphics displays using a layout widely employed among decision makers. The interface was computer generated hardcopy. By Anthony's criterion (Mitroff and Mason, 1977), the organizational setting was one of managerial control of resources. These components were the key factors behind the design of the experiment.

Chapter III, Methodology, describes the experiment in detail. It discusses the sampled population of Arizona State University (ASU) students and their characteristic factors of age, sex, college level and experience, attitude toward the computer, experience with computers and video games which were of interest to this research. It discusses and justifies the use of the selected inventory case study employed as the instrument for data collection. The mixed graphics display or mode of presentation is then described and presented in Figure 3.1, with the four color alternatives and one monochrome intensity pattern defined. The reasons for choosing the hardcopy interface are enumerated. The organizational context of managerial control is considered.

This chapter also identifies the performance measures

of time and quality of decision, the two most coveted resources of decision makers. These two are subdivided into 15 categories, 3 which are a function of time and 12 which are a function of quality of decision. Each type is operationally defined as it pertains to this research.

Pretesting of the instrument and the experimental design are presented. The procedures and results are given. The Methodology chapter concludes with the experimental procedures whose objective was to collect data, which, when analyzed properly would lend insight into how the attribute of color affects the performance of a decision maker.

Analysis and Results (Chapter IV)

Analysis was done in three areas: the problem statement, the data and the statistics which were generated from the data. Chapter IV, Analysis and Results presents all three in detail.

Initially the problem statement was analyzed. Using 15 measures of performance and 7 characteristic factors of the sampled population, the problem statement was subdivided and formalized into testable hypotheses (Table 4.1, Research Formal Hypotheses) which allowed for rigorous and indepth analyses.

The data was systematically analyzed, using the

schematic shown in Figure 4.1, Analysis Strategy. This was done to insure the 2-factor fixed effects ANOVA model used for analyses of the research question was not only appropriate based on the hypothesis but also on the data. Special attention is given to this analysis as rigorous analytic procedure is one major shortcoming noted in the literature.

The Analysis and Results chapter concludes with the results and conclusions derived from analyzing the statistics of the appropriate mathematical model. Of the 111 hypotheses that were tested, the data underlying the majority of them (102) contained insufficient evidence to conclude that any of the five treatments (4 color, 1 black/white) had a significant differential effect on the performance of a decision maker. Even after the effects of the decision maker's characteristics such as age, sex, college level and experience, attitude toward the computer, experience with computers and video games were taken into account in the variation of the mean performance, the color attribute did not test statistically significant at the .05 significance level.

For the nine hypotheses whose data contained evidence that supported the contention that color in a hardcopy mixed graphics display did affect the performance of

decision makers, in-depth analyses were accomplished. Results of these studies are displayed in Table 4.15, Ranking of Treatments. The trends among these results showed that the top ranking treatment was consistently chromatic, though not always the same particular set of colors. However, the colors presently preset on a majority of CRTs were consistently above average relative to the users' decision quality performance. However, the monochrome intensity pattern always rated second, and never significantly different (α=.05) from the chromatic treatment ranking first. The seven cases where the CRT colors outscored the monochrome, the average increase in performance was approximately 9.2 per cent. The two cases where the monochrome outscored the CRT color, the relative decrease in score using the color was about the same, 7.0 per cent.

Conclusions and Recommendations (Chapter V)

Conclusions, summaries and recommendations are detailed in Chapter V. Conclusions and summaries categorized initially by the particular performance variable, then by performance variable type (decision quality and time) and finally with respect to the general problem statement are offered. This final conclusion, which is the overall contribution of this research, is

that the attribute of color, when integrated functionally in computer generated mixed graphics displays does not consistently affect the performance of a decision maker who has an engineering background who solves structured problems based on deterministic data presented within a managerial role of control of resources. When performance was affected in the positive direction, the increased effectiveness relative to monochrome was always statistically insignificant with the practical difference minimal. If color is integrated into the mixed graphics display, the typically used CRT alternative appears to be the best, however a decision maker should not assume it will significantly increase performance.

This chapter and research concludes with recommendations for further research. These recommendations are based on the recognized limitations within this research, and the suggestions and voids in the MIS literature.

II. LITERATURE SEARCH

Introduction

When organizational resources are expended for incorporating the attribute of color into an existing or developing graphics decision support system (DSS), what benefits does the organization hope to gain in return? The literature has just begun to address this question. However the research on the effects of using color in visual displays in other tasks has been going on for years by noted Human Factor specialists and others. These reports and findings can be categorized under one of the three headings: theoretical support, subjective reasoning, and objective and experimental findings. This literature review is developed according to those categories. The findings are then summarized and compared, with voids in the present knowledge highlighted. Emphasis is then given to the primary gap that exists in the knowledge available for managers who desire to evaluate the benefits of color visual displays to their organizational effectiveness.

Literature Findings

Introduction

The literature search utilized the Arizona State

University computerized literature search service. The data bases queried included INSPEC, NTIS, COMPENDEX, ERIC, Dissertation Abstracts and Microcomputer Index. Abstracts from the queries were reviewed to determine the source documents appropriateness. Potentially useful papers were located and read. Relevant references from these reports were also examined. Where possible the source document for each study found particularly key to this research was found and critically reviewed. In addition personal communications were established between the researcher and several noted authors in the Human Factors field of color research: Richard Christ, Thomas Tullis and H. Rudy Ramsey. Their personal critically annotated bibliographies were provided to complete the literature search.

These efforts yielded findings in the relevant topics of the use of color in information systems and the need for research in the area of DSS that incorporate the option of color. What follows is the review of the relevant literature on the use of color as an information code in visual displays. It is developed by sections: theoretical support, subjective reasoning and experimental findings.

Theoretical Support

> The integration of color into the visual imaging process expands our capability of visualizing complex mental constructs and communicating sophisticated ideas (Truckenbroad, 1981, p. 86).

During the processing of information that is linearly sequential, read left to right and from top to bottom, the left hemisphere of the brain is activated. When the attribute of color is added, the right cerebral hemisphere is evoked to process the information content simultaneously with the left side. Based on this biological concept, many predict information presented in color will be processed much more efficiently than that which is not (DeMars, 1975; Miller, 1982; Truckenbroad, 1981). What color(s) are best, how much more efficient the process is and what other variables impact this processing is just beginning to be explored in information systems.

Subjective Reasoning

The decision to use color is often based on subjective feelings rather than scientific findings (Karner, 1975). Until very recently it was relatively difficult to investigate the variable of color on computer based systems simply because of the cost associated with the system (Durrett, 1979). However people have grown up

with the color media and have come to expect this attribute of color in other technology including their graphic displays. Consumers are justifying their use of color systems based on their and others' subjective feelings.

DATAMEDIA, which introduced its colorscan color CRT in 1981, feels it is the brilliancy and clarity of the color that sells the product (Miller, 1982). Visual packaging is critical to merchandising and color is an integral part of any high quality graphics package (Catalano, 1982; Judd and Wyszecki, 1975, Needle, 1982). People are buying and using color graphics packages as part of their DSS.

When asked why color is preferred, several common responses occur across different populations of users. Some feel color adds another dimension to the graphics output (Kirby, 1979). Others contend it significantly increases productivity though they have not scientifically determined if in fact color has this effect (Friend, 1980; Miller 1982). Another argument for color is that it is easier on the eyes (Christ, 1977; Miller, 1982). Many justify the use of color through arguments of prestige and motivation but have not yet rigorously analyzed this latter hypothesis (Belie and Rapagnani, 1981; Durrett,

1979; Eddy, 1979; Waller, LeFrere and McDonald-Ross, 1982). Eddy based this comment on a study he did in which 24 students used color display terminals in a biology laboratory class. No control group or monochrome terminals were used. Durrett (1979) felt many students who previously were not interested in computers were attracted to them based solely on the color display system's intrinsic novelty.

Some purport color adds more information per character on the display, therefore making color displays more cost efficient than their monochrome counterparts (Miller, 1982; Morris, 1979). Skellings, who has a patent pending on electronic displayed color poetry and is widely published in electronic communication periodicals, feels he gets "something like 45 percent more information" (Miller, 1982, p. 87) on a color display. He did not attempt to verify his statement. Another claim presented without research backing yet is commonly accepted is that information presented in color is recalled better (Durrett et al., 1981). Another reason given is that color is more natural thus making graphs easier and more natural for the eyes to follow (Eddy, 1979). The most prevalent response to the question is that color is highly appreciated and aesthetically pleasing to the user (Bergman, 1980;

Carlson, Parent and Csuri, 1979; Catalano, 1982; Durrett et al., 1981; Eddy, 1979, Tullis, 1981).

In several experimental studies, subjective opinions were solicited which were not always backed by the data findings. In one study pilots were exposed to two types of visual display systems (black/white and color) during landing approach exercises. They favored the color displays. They felt color resulted in greater pilot relaxation and decreased fatigue. They also felt color gave the display a higher picture quality and allowed for greater real depth perception (Chase, 1970). Subjective evaluations were the emphasis of this study.

Another study which solicited subjective responses involved subjects who took part in a simulated military tracking experiment. They reported feeling more secure when using a color coded display. They commented that "color helps...(you) can see the overall pattern better...color makes it (task) easier" (Munns, 1968, p. 1221). The objective findings were not concurrent with the subjective comments.

Color is becoming a widely demanded attribute of many information display systems. Unfortunately the decision and justification to incorporate this attribute is more often than not based on the subjective feeling that "it

couldn't hurt" (Karner, 1975, p. 227) and "color graphics are so obviously much better than black and white and they're not really that much more expensive anymore" (Miller, 1982, p. 92).

Experimental Findings

Investigating the effects of color has fascinated researchers for over 25 years and continues to do so with the advent of the computer technology that now supports affordable color and the increasing reliance on these computers and their outputs. A majority of the historical studies deal with the impact color has on time and error rate when the human subject is involved in a simple task. These tasks included locating, searching, identifying, comparing, counting, tracking and recall (Barker and Krebs, 1977; DeMars, 1975; Keister, 1981). For all of the studies color was one of the two or more codes under investigation. Code is defined as "any system of symbols for meaningful communication" (Cook, 1974, p. 3). Rear projection slides, viewgraphs, hardcopy and/or CRT terminals were the common display devices used, with the first two being the most prevalent. A display device is the "principle interface between the observer and the computer or other device generating the information" (DeMars, 1975, p. 2-1). All studies reviewed were done

under laboratory conditions. Subjects for the most part were not experienced in the tasks they undertook.

The results of the historical experiments can be summarized by stating the relative effectiveness of color is a function of the task undertaken (Baker and Krebs, 1977; Christ, 1977; Christner and Ray, 1961; Cook, 1974; Tullis, 1981).

The studies will be discussed in chronological order beginning with those done in the early 1950s. Where available the purpose of the research, the key variables, the specific task, the sample subjects and the conclusions are delineated. A summary by researcher (Table 2.1) and by task (Table 2.2) conclude this section complimented by a comprehensive analysis by Christ, Corso and Teichner of the studies to date.

Green, McGill and Jenkins, 1953 were interested in variables that affected response time in a simple search task. Initially they determined the most important variable is the density of the objects to be scanned. Using color code of two categories (blue and yellow), they found search time was significantly decreased when color was used as a partially redundant code (Barker and Krebs, 1977; Cook, 1974; Jones, 1962).

Green and Anderson, 1956, used color as a partially

redundant code in a search task. Search time was significantly decreased under the condition that the observer knew the color code prior to engaging in the task (Barker and Krebs, 1977; Cook, 1974).

Christner and Ray, 1961, wanted to determine the relative effectiveness on performance of selected target background coding combinations when used in locating, counting and identifying tasks. The coding variables were numerals, shapes and colors. They concluded that the choice of target codes was task dependent. Color was superior, though not significantly, when used in locating and counting tasks. It was not for identifying tasks. Subjective preference was for color code across all tasks (Christner and Ray, 1961).

Hitt, 1961, analyzed the relative effectiveness of using selected abstract coding methods on operator performance when engaged in various tasks. Operator performance was measured using number of correct responses and speed of response per experimental condition. Two sets of 5 subjects were instructed to complete five map reading tasks (identify, locate, count, compare and verify) as rapidly and accurately as possible. The codes were color, numerals, letters and shapes. To investigate both performance measures, time and accuracy were combined

in the form of number of correct responses/minute. Based on an analysis of variance of a fixed effect model on the mean correct response/minute scores, Hitt concluded that color and numerals were the two superior coding methods in this reading situation. He concurred with Christner and Ray that numerical coding was more effective in identification type tasks while color was more effective for the locate, count, compare and verify tasks (Hitt, 1961; Jones, 1962).

Newman and Davis, 1961, concluded that color coding reduced response time in both locating and decoding tasks (Cook, 1974).

Schutz, 1961, studied the effect of color and black/white coding when subjects were engaged in point reading and comparing tasks. Red, yellow, green and purple were the colors used. The first task consisted of reading vertical-axis values from horizontal axis values for a given coded line. The comparing task consisted of determining the coded line with the highest value on the vertical axis for specified values on the horizontal axis. Performance of ten subjects was measured as time to complete the task. Errors were few and random. Based on his analysis of variance on the mean performance time scores, Schutz concluded the color only slightly improved

the rate for point reading but not for comparing (Schutz, 1961).

Jones, 1962, recommended the task be the criteria upon which the decision be based whether to use color coding. Results from his analysis suggested color codes are not suited for tasks that demand rapid and precise identification but are valuable in decreasing search time in a locating task (Barker and Krebs, 1977).

Smith published his research on the effects of color on speed and accuracy over various tasks from 1962 through 1965. His experimental designs and dependent variables were similar across all his studies.

Smith, 1962, had 12 subjects do 300 visual searches on 12 inch square slides that were placed 18 inches away. His analysis concluded color resulted in faster and more accurate searches when the target color was known beforehand. Neither the particular color of target or background had any significant effect on performance (Barker and Krebs, 1977; Cook, 1974; Smith, 1962).

Smith, 1963, concluded that color coding resulted in 65% reduction in time for visual search, 69% reduction in time for counting and 76% reduction in counting errors. His findings were based on data collected on 12 subjects (Barker and Krebs, 1977; Cook, 1974).

Smith and Thomas, 1964, had 8 subjects engage in a counting task. The display device was a 29 inch square which was placed five feet away from the subject. They concluded color coding resulted in faster performance times and fewer errors (Barker and Krebs, 1977; Cook, 1974).

Smith, Farquhar and Thomas, 1965, designed an experiment that assessed and compared the effects of symbolic, numeric and color coding in formatted displays. Row comparison and item counting were the tasks performed by their 12 subjects. Based on analysis of their mean time scores and errors they concluded that color coding was significantly superior to any other coding. Color coding reduced performance time over no coding by 72% for the item counting task and 47% for the row comparison task. Their overall generalization was the effect of color depends upon the joint interaction of the format of the display and the task to which it is applied (Barker and Krebs, 1977; Cook, 1974; Smith, Farquhar and Thomas, 1965).

Hoisman, Hannah, and Schaef, 1963, were under contract with the Air Force (AF) to provide an experimental investigation of visual displays which would lead to the specification of information presentation

parameters associated with the display of processed information to AF intelligence officers. After completing a thorough search of the literature Hoisman et al discovered there was a gap in knowledge concerning visual display presentations and their associated parameters. They designed a project to gain information that would narrow this gap. This project was designed in the context of the tasks, presentations, parameters and mission of the AF Intelligence System.

Initially Hoisman et al determined through factor analysis of data from surveys and interviews the display type, the tasks and parameters most representative of the Intelligence community. Map displays were the most utilized. Six tasks were identified when using a map: locate (i.e. How many cities?), identify (i.e. Is this target an air base?), extract (i.e. What is the city's population?), integrate (i.e. What is the target value of this city?), correlate (.e. Is the strength of these two bases equal?) and analyze (i.e. What are the rank order of the targets?). Density and color coding were determined to be the major display parameters as they appeared to contribute a large portion of the within display differences in performance.

Using hardcopy maps as the display device,

experiments were developed to investigate gross qualitative indices of density and the effects of color coding on subject performance when completing the six categories of tasks. Four tentative hypotheses were proposed, with one of particular interest to this study: Color coding enhances performance.

A pilot study was conducted to pretest the experiment. Volunteers were 43 males of freshman or sophomore academic standing enrolled in a psychology course at the University of Pittsburgh. They were paid a minimum hourly wage. These males, ranging from 18 to 20 years old, all passed the Dvorine Psuedo-Isochromatic test for color vision prior to the experiment. After a 20 minute lecture familiarizing the subjects with the task, and a practice session, each was given an envelope containing maps with a legend and a list of tasks (questions) to be completed. Responses to the questions, all which had answers, were made by writing on paper visible through an open window of a mechanical instrument located in front of them. This instrument insured no answer could be changed. It also recorded the time per task completion. The maps used were generic maps utilized by the AF, altered by removing the geographic names and replacing them with randomized letter codes. This

controlled for any familiarization of the maps by the subjects. The effects of order presentation were considered negligible since the materials and the tasks were distributed within each package in random order. The effects of location of significant data on the maps were controlled by completely randomizing the location points. The dependent variables or measures of performance were speed and accuracy. Speed was defined as the mean response time per task (question). Accuracy was measured as total number of correct tasks completed. The independent variables were color, contrast, task and task level.

Using a two factor experimental design and analysis of variance, Hoisman et al reported their data supported their hypothesis that color tended to reduce the effect of high density. Correlation analysis of the data suggested no significant correlation between speed and accuracy.

Based on their findings from their pilot study, Hoisman et al performed their main experiment to evaluate the effects of color, density and task on the performance of AF intelligence analysts using trend graphic, cartographic and tabular displays. The volunteers were 68 male college students who were paid a minimum hours wage to participate in five 3-hour sessions. They were given

the incentive of being paid extra if chey performed better than par.

The experiment was given in a windowless classroom illuminated by fluorescent tubes to 40 footcandles. Each worked at a 24x48 inch flat top 28 inch high desk. As in the pilot study a pretest training and practice session was given. Each subject was then given a package containing tasks (3 simple, 3 complex questions) and maps. Each worked at his own pace.

The independent variables were density (two categories of 5 levels each), color (chromatic (red, blue, purple, shaded grey) and black/white) and tasks (6 types previously defined). The dependent variables were accuracy and speed item by item and in total.

The underlying assumptions stated were the following. Their subjects did not differ significantly in any of their essential characteristics from AF intelligence analysts as a group. The testing materials used and the principles followed in their design were meaningfully related to the actual materials used by the analysts. Finally the tasks used in testing the hypotheses and the methods used in deriving them related to the tasks performed by the analysts.

Using a complete 5x5x2x6 factorial design, a standard

analysis of variance technique was employed to examine both the major and interactive effects of each of the 18 variables. Standard product moment correlation analysis indicated a nonsignificant relationship between accuracy and speed.

Hoisman et al's major conclusions were several. Of prime interest to this research is their stated conclusion the

> the addition of a color as an enhancement of a symbol's distinguishability does not increase performance and may even lessen it (Hoisman et al., 1963, p. 7)

Color here was added coding (redundant) that was used with, not instead of, shape and size codes.

Brooks, 1965, studied the effects of using different redundant color codes in a search task. He varied the colors and the number of colors used. Coding consisted of adding a bar of color beneath the target. His conclusions were based on data collected from six groups of 10 subjects each. All subjects were informed in advance their particular color coding condition. Color resulted in significantly better search times than the no color condition. Neither the number of colors, the particular color, the display background used, nor any interaction of these variables had any significant affect on performance time (Barker and Krebs, 1977; Brooks, 1965; Cook, 1974).

Munns, 1967, was interested in the effects of display symbol variation upon operator performance in aircraft interception. Two levels of density (high and low) displays and color (monochrome and two color (red and blue)) coding of symbols were used as independent variables. All combinations were given to each of 12 Lehigh University male subjects. Their task was to detect and discriminate enemy and friendly aircraft presented on an 8 x 10 problem sheet inserted in a box that was used to simulate a radar display. Every effort was made to reduce the subject's decision process to the simplest level. Time to complete the task and number of errors (enemy aircraft missed) were Munn's dependent variables. Using the results from a 2 way analysis of variance he concluded that color reduced performance time and made the subjects feel more secure but there was no indication color reduced subject errors. Display density or the interaction between density and color were not significant (Munns, 1968).

Lee, 1969, assessed the comparative effectiveness of colors in a target search task. Thirty-six subjects were timed on their search and count tasks of targets displayed on a color CRT. Color was superior to black/white coding for both tasks (Barker and Krebs, 1977).

Dooley and Harkins, 1970, designed an experiment with the purpose of examining the effect of color on attention and recall. A standard flipchart was their display device. On it color (treatment) was used in three ways: functional, attention-getting and monochrome. Functional color was defined as employing color as the primary information code. If the color was removed the communication was degraded. Attention-getting color was defined as using color to attract attention. If the color was removed the communication was not degraded. Monochrome color was defined as the void of color or the use of black/white shadings as the communication medium. Dooley and Harkins' performance measures were scores derived from a post treatment written test on recall, and the number and amount of time the subjects looked at the flipchart. As far as the subjects knew the flipchart was irrelevant to the experiment in which they were engaged. It just happened to be standing in the corner of the experimental room.

Twenty-three men and 22 women employees of Xerox Corporation were randomly assigned to one of the treatment conditions. The flipchart contained a 71 x 97 cm chart displaying bar graphs of comparative information of three companies on three different topics. Xerox was one of the

companies. The legend and bar graphs were coded as monochrome (black/white/shading), functionally colored (red green blue) or attention getting colored (monochrome configuration with colored borders). When the subject entered the room to participate in another study the treatment was mounted on a flipchart stand in a corner approximately 30 degrees to one side of the subject. It was not mentioned by the experimenter who was evasive if asked about it by the subject. The experimenter left the room on two occasions for 2.5 and 1 minute respectively for valid reasons.

Dooley and Harkins used a camera to record number and length of times subject spent looking at the flipchart. These data were then order ranked for all participants. Measure of attention was the sum of the rankings of the two recorded indexes. A Kruskal-Wallis test on the mean attention ranks showed the three treatments to be different though a subsequent chi-square test indicated a significant difference only between color and monochrome. From this Dooley and Harkins concluded more attention was given to the colored charts.

At the conclusion of the experiment the flipchart was covered and each subject was asked to complete a written test whose questions were based on the flipchart's

information. The questions ranged from general to specific, each with a score of one point if answered correctly. Measure of recall was the sum score. The subjects' overall scores were then ranked and analyzed using a Kruskal-Wallis one-way analysis of variance test. There was no evidence to support the hypothesis that color had an effect on recall. From this they concluded

> it is improper to assume that the addition of color to graphic communications will increase their effectiveness whether the color is used functionally or decoratively...There are undoubtedly applications in graphic communications where the addition of color would have a clear beneficial effect over that expected for the best available black and white code, but the present example is not one of those cases. Since information on the effectiveness of color in graphic communications has substantial practical significance, more research in this area is clearly needed (Dooley and Harkins, 1970, p. 854).

Beyer, Schenk and Zietlow, 1971, compared the effectiveness of color coding to that of brightness coding when using the CRT as the display device. Effectiveness was measured as task response time and subject preference. Sixty-eight subjects were involved in four experimental tasks: locating, tracking, identifying and giving subjective evaluation. Using statistical analysis techniques, response time was significantly faster for the color coding for the first three tasks. However Beyer et al. comment the difference was probably not of practical

importance. Subjective preference favored the color coded displays (Barker and Krebs, 1975).

Kanarick and Petersen, 1971, recognized that few studies had examined the effects of redundant coding, particularly color, on retention when the pay off ratios were varied. Payoff ratios x:y were defined as the number of points awarded for a correct response/column. For half of the columns x points were scored for a correct response. For the other half, y points were scored. The subjects were informed prior to the experiment the point value of each column. Using a mixed factorial design with the coding condition being the between subjects variable, they analyzed the effect of using or adding color to a code. The task used was one of remembering how many times a particular code was presented in a particular column. They had ten columns (A-J), and six codes. The codes were three types: numbers (1-6), colors (blue, violet, yellow, red, orange, green) or uniquely colored numbers (blue 1, violet 2, etc.). Three treatment sessions were administered: training, 4:1 and 1:1. The training session familiarized each of the 56 paid undergraduate students with the task. During the 4:1 payoff session, the subjects were told they would receive four points for a correct response in specified columns and one point for a

correct response in the other half. The third session was the same as the second except one point was given to all correct responses. Subjects competed for monetary bonuses based on total points. Per cent correct response was the measure of performance.

Based on analysis of variance and Duncan's New Multiple-Range Test, Kanarick and Petersen concluded that redundant color coding of numerical information did not facilitate performance. Subjects were more attentive to the point value of the column and not the color when the payoff was 4:1. Redundant coloring did not enhance performance with the 1:1 payoff condition (Kanarick and Petersen, 1971).

Shontz, Trumm and Williams, 1971, chose aeronautical charts as the display device in their examination of the effects of color on performance when searching for geographic checkpoints. The ground/map coordination task was similar to that of a pilot who had to locate his position after being geographically disoriented. Cumulative search time was the measure of performance. The independent variables of interest were coding (uncoded chromatic, coded chromatic, coded achromatic), clutter and code size.

To gather their data, Shontz et al. had 33 junior and

senior Air Force Reserve Officer Training Cadets who were currently enrolled in pilot and navigator training locate 48 checkpoints on a 14 x 14 inch map. They were seated in a modified tilted dental chair at a viewing distance of 20 inches. Shontz et al. identified the 28 color codes in Munsell Book Notation. A mixed design was used with map coding as between-subject variable and background clutter and code-alphabet size as within subject variables.

Results from the Kolmogorov-Smirnov two sample test indicated differences between distributions of cumulative search times was significant between those using color coded maps and uncoded maps. Location of checkpoints was faster for those having color coded maps. However the difference was not significant when comparing coded chromatic and coded achromatic conditions, despite the fact those viewing achromatic maps reported difficulty distinguishing between roads and rivers of similar configuration. Based on this finding Shontz et al. question whether properly designed and coded achromatic maps would improve performance sufficiently under varying conditions to warrant their consideration (Shontz, Trumm and Williams, 1971).

Markoff, 1972, varied the resolution in his study on the impact color had on target recognition. Analysis of

the data on 48 subjects indicated that chromatic real world imagery resulted in significantly faster target identification times than did comparable achromatic screens. This effect was more pronounced as the resolution decreased (Barker and Krebs, 1977).

Wedell and Alden, 1973, tested the hypothesis that color was superior to numeric coding in a modified keeping track task particularly as the total number of items displayed is increased. Accuracy was the measure of performance for the 36 male subjects who simulated the task of an air traffic controller who had to detect changes from display to display. The colors used were blue, purple, red, orange, yellow and green. Based on analysis of variance analysis on error type they concluded color can aid in retaining information concerning number of items presented and their location but does not help in identification (Wedell and Alden, 1973).

Saenz and Riche, 1974, conducted a study that examined the effects on search time when shape and color codes were used separately and redundantly. Target and backgrounds were varied by color and shape. The three independent variables were color code, shape code and color shape (redundant code). The colors were identified by Colormatch Paper and include red, yellow-red, yellow,

green-yellow, green, blue-green and blue. The shapes used were circle, hexagon, triangle, cross, diamond, star and square. The dependent variable was time required for each of the 24 male undergraduate students to locate and respond to six targets in each of 48 settings. The display device was a thick white illustration board that contained the target and background stimuli which was inserted in a tilted box for ease of visual access. All students were required to pass the Dvorine Pseudo-Isochromatic test of color vision.

Saenz and Riche performed an analysis of variance on the logarithmic transform of search time in seconds and applied Tukey's HSD test for pairwise comparisons among the means of main effects. Redundant and color codes were found to be equally effective and more so than shape code. The authors suggested that although redundancy may not degrade operator's response, it may not help it either. Therefore cost should be a factor in the design of the code for display (Saenz and Riche, 1974).

Cahill and Carter, 1976, researched the impact of display density and color code size in a search task. Twenty subjects were asked to search for three digit numbers on displays ranging in density from 10 to 50 items coded in 1 through 10 colors. Based on their analysis of

search time, they recommend seven colors as the optimal number of colors to be employed. Fewer and more colors increased search time (Barker and Krebs, 1977).

Keister, 1981, was one of the first to investigate and report on the effects of color coding in other than the traditional search, locate and/or identify type tasks that usually concentrated on random patterns of letters, numbers, geometric shapes and color. Keister's contribution involved the organized task of data entry at two levels of complexity: adding and changing an inventory item. Time/transaction and errors were the measures of performance. Two laboratory experiments were designed to ascertain the effects of color coding on operator performance when entering or changing product information in an inventory data base. The mode of visual display was a color or green phosphorus CRT with keyboard interface. Data format was color or monochrome tabular text. All subjects were experienced in typing and data entry but not in the specific task.

The first experiment involved having the eight paid female subjects enter data in two sessions and change it in two other sessions. Each subject used both codes (monochrome and color) for each task, presented in random order. The authors reported the overall results showed

data entry of a new item was somewhat more rapid when using the color code. The data indicated that color accounted for even greater speed for the change task. The increase was about two seconds/entry which was not statistically significant. After analyzing the order of presentation the authors suggested the data supported the conclusion that color facilitates learning and offers a permanent advantage in speed for those who continually switch data entry tasks.

In their second experiment 54 National Cash Register (NCR) software development personnel were asked to complete the data change task as quickly and accurately as possible. Green phosphorus, compatible color and conflicting color presentations were used by the subjects during their 24 transactions. Color usage was compatible across tasks if a particular color always symbolized the same information (i.e. yellow always denoted price). Color usage was conflicting if a particular color symbolized different information across tasks (i.e. yellow denoted price in one task and inventory quantity in another). Density was varied by number of products listed on screen: one, two or four. As in the their first experiment, the overall effect of color on total time was not significant. However the authors suggested that color

was superior when the subject was using the most complex screen which contained four products.

Keister concluded the overall performance on color displays was superior to that on monochrome displays despite the lack of statistical significance. Based on his reports that color was superior in all experiments, Keister suggested a motivational interpretation. Color can increase operator motivation and facilitate performance by making the task more interesting. In moderately complex data entry task (locating an item and keeping track of entry activities) color provided only a minimal advantage. In a more complex task (which required the entry of a number before making a change), color aided in the locating and keeping-track tasks. Color also facilitated the switching among different type tasks. These results were based on mean times of performance that were not statistically significant. The author did not indicate how subjective opinion was obtained from the participants though he purports color added to motivation and decreased boredom.

Tullis, 1981, also designed a study that used other than the traditional search, locate and/or identify type tasks. He described his task of diagnosing the condition of telephone lines as a typical example of a large class

of computer-aided decision-making tasks. Subjects had to integrate various classes of data in order to make a timely decision about the status of some real world system. His setting was a day training session for eight Bell Laboratory employees who had experience in the Bell System ranging from 4 to 27 years though not in the particular tasks used in the experiment. The display devices were four types of formats presented on a color graphic CRT: narrative (complete words, and phrases), structured (fixed tabular form), black and white graphics (structured plus schematic and shape coding for condition of telephone lines), and color graphics (black and white graphics format with color coding substituted for shape coding for condition). The colors used were consistent with the population stereotypes: green for good, red for bad and yellow for marginal. The day was divided into 10 training sessions. The first six sessions were text-based that familiarized the subjects with the electrical characteristics of telephone lines. The last four sessions were CRT based and were designed to acquaint the subject with one format/session and then test the subject using two sets of 37 different displays accompanied by a 24 question multiple choice test. Subjects completed a format session after they achieved at least 80% accuracy.

They then proceeded on to the next session. The questions ranged from simple identification to complex integration and decision making. Speed, accuracy and number of training sessions needed were the measures of performance. A subjective questionnaire evaluating the four formats was administered at the day's end.

To control for practice effects Tullis used a 4 x 4 Latin Square design to determine the order of presentation of format/subject. Analysis of variance of data indicated that accuracy did not vary with format. The 80% training accuracy criterion may have eliminated any significance. However response time was affected by format. Comparisons of means using Tukey's HSD test suggested the response time when using the graphics format was significantly shorter than when using the narrative. There were no significant differences in response times between the two graphics conditions of color and black/white. Fewer training sessions were required to achieve the required accuracy for the graphics over the narrative format.

The only difference between the two graphics formats arose in the subjective ratings. Seven of the eight subjects indicated they preferred color over black/white and found the color format aesthetically pleasing (Tullis, 1981).

Boeing, 1982, completed a study whose purpose was to evaluate the CRT pictorial formats and their effectiveness in presenting integrated information to the pilot (Scott, 1983). One of the issues under investigation was how the color formats compared to those in monochrome. Their analysis was based on data collected on 18 experienced military pilots who flew two 20 minute simulated air-to-air and air-to-ground composite missions. Their cockpits were outfitted with CRT instruments. One of the display configurations was monochrome, similar to the current generation aircraft CRT display. A second was a color vector graphics display which produced stick-like figures. The third configuration used a color raster scan system which created solid color drawings.

The subjective and objective results differed. The pilots overwhelmingly preferred the color over the monochrome configuration. However, the objective data on performance, measured as speed in interpreting and responding to display information, did not contain sufficient evidence to report any differences between the three configurations. Mr. Reising, from the Flight Dynamics Laboratory, commented the study did show

> that color pictorial displays probably ease the pilot's information assimilation task when viewing a complex display...although more studies are required (Scott, 1983, p. 232).

Christ, Corso and Teichner, 1975, 1977 and 1983 are presently the key figures in the field of study of the effects of color on human performance when used in a display for information transfer. All of their studies here were supported under contract by the Office of Naval Research with the purpose of evaluating the basis of possible design recommendations for or against the use of color in aircraft displays. Their first step was a detailed search of the literature and analysis of findings (Christ, 1975). Based on this analysis and noted glaring gaps in knowledge, experiments were designed and carried through in attempts to narrow some of these gaps. Experience of operator and level of task were their first concerns. Previous studies had unskilled subjects doing simple tasks. Christ and Corso (1975) trained their subjects for almost a year on tasks of varying difficulty before collecting data. In 1977 Christ, Corso and Teichner expanded their study to include experienced operator performance in a more real world type environment. Christ had the results of all their studies published in 1983 with the conclusion that

> the relative effectiveness of different visual display codes depends upon other display conditions, the task, and the dependent measure used to make the comparison (Christ and Corso, 1983, p. 83).

This paper's literature review concludes with a summary table of the experimental research previously discussed (Table 2.1, Summary of Experimental Research on the Effects of Color) and Christ et al.'s works. These authors' contributions serve as a summary not only because of their recency but because of the contents. Their thorough review and analysis of the literature took into account all studies reported above except six. The works of Hoisman et al., 1963, Lee, 1969, Beyer et al., 1971, Cahill and Carter, 1976 and Keister, 1981, Boeing, 1982 were not included either because of data requirements or date of publication.

Table 2.1, Summary of Experimental Research on the Effects of Color by Author and Date, presents a concise skeleton outline of the works reviewed in this experimental section. As in the narrative, the researchers are listed in chronological order. The tasks and sample size from which their results were generalized are given where available. The dependent variable or performance measure is stated. The Color Significant entry indicates whether their results showed that using

Table 2.1

SUMMARY
EXPERIMENTAL RESEARCH
on the
EFFECTS of COLOR in VISUAL DISPLAYS
by
AUTHOR and DATE

RESEARCHER(S)	DATE	TASK TYPE	No. Sub.	PERFORMANCE MEASURE	COLOR SIGNIF
Green et al.	1953	Search		Time	Y
Green Anderson	1956	Search		Time	Y
Christner	1961	Locate	5	Time	Y*
Ray		Count	5	Time	Y*
		Identify	5	Time	N
Hitt	1961	Locate	10	Accuracy/Time	Y
		Count	10	Accuracy/Time	Y
		Identify	10	Accuracy/Time	N
		Compare	10	Accuracy/Time	Y
		Verify	10	Accuracy/Time	Y
Newman	1961	Locate		Time	Y
Davis		Decode		Time	Y
Schutz	1961	Pt. Read	10	Time	Y
		Compare	10	Time	N
Jones	1962	Identify		Time	N
		Locate		Time	Y
Smith	1962	Search	12	Time	Y
			12	Accuracy	Y
Smith	1963	Search	12	Time	Y
		Count	12	Time	Y
			12	Accuracy	Y
Smith	1964	Count	8	Time	Y
Thomas			8	Accuracy	Y
Smith Thomas	1965	Count	12	Time	Y
Farquhar		Compare	12	Time	Y
Hoisman et al.	1963	Identify	68	Time	N
			68	Accuracy	N
		Locate	68	Time	N
			68	Accuracy	N
		Extract	68	Time	N
			68	Accuracy	N
		Integrate	68	Time	N
			68	Accuracy	N

		Correlate	68	Time	N
			68	Accuracy	N
		Analyze	68	Time	N
			68	Accuracy	N
Brooks	1965	Search	60	Time	Y
Munns	1967	Srch/Idnt	12	Time	Y
			12	Accuracy	N
			12	Subjective	Y
Lee	1969	Search	36	Time	Y
		Count	36	Time	Y
Dooley	1970	Attention	45	Time	Y
Harkins		Recall	45	Accuracy	N
Beyer	1971	Locate	68	Time	Y
Schonk		Identify	68	Time	Y
Zietlow		Track	68	Time	Y
Kanarick	1971	Recall 4:1	56	Accuracy	N
Peterson		Recall 1:1	56	Accuracy	N
Shontz et al.	1971	Locate	33	Time	N
Markoff	1972	Identify	48	Time	Y
Wedell	1973	Recall	36	Accuracy	Y
Alden		Identify	36	Accuracy	N
Saenz Riche	1974	Locate	24	Time	Y
Cahill Carter	1976	Search	20	Time/Density	Y
Keister	1981	Data Entry	8	Time	Y*
		Data Entry	54	Time	Y*
Tullis	1981	Diagnostic	8	Time	N
			8	Subjective	Y
Boeing	1982	Inter/Resp	18	Time	N
			18	Subjective	Y
Christ	1975	Identify	8	Time	N
Corso		Srch/Locate	8	Time	Y
		Recall	8	Time	N
Teichner	1977	Combined	8	Time	N
Christ		Combined	8	Time	N
Corso			8	Accuracy	N
		Integrate	8	Time	N

Y* not statistically significant

color as a code improved performance over other achromatic codes. Statistical significance was the criterion. The method of analysis is not shown as the majority of the studies reported applying the fixed or mixed effects analysis of variance techniques. However it was noted that the assumptions validating the use of the mathematical model were seldom mentioned or addressed in any of the studies' publications.

Christ, 1975, searched the literature from 1950 through 1973 for studies involving experimental research on the effects of color when used as a code for information in visual displays. Forty-two studies were found that satisfied the following five criteria. The dependent measure was per cent correct, information transmitted or response time. Second, the data yielded a comparison between chromatic and achromatic codes and were available. Third, the task required information extraction from symbolic or pictorial displays rather than from direct viewing. Fourth, subjects were normal young adults. Finally, the procedures and general experimental design were acceptable.

In order to analyze the data among different studies which used different procedures for measuring the same dependent variables, Christ normalized the scores within

each experiment while holding the other task parameters constant. For the dependent variable accuracy, the score was calculated using

$$\% \text{ Difference in Score} = \frac{\text{Color} - \text{Achromatic}}{\text{Achromatic}} \times 100 \qquad (1)$$

Search time was calculated using

$$\% \text{ Difference in Time} = \frac{\text{Achromatic} - \text{Color}}{\text{Achromatic}} \times 100 \qquad (2)$$

Positive values for either score indicate an advantage for color, negative scores a disadvantage for color, relative to a particular achromatic target attribute.

Christ categorized the results of the acceptable studies by the way color was used on a target (unidimensional, multidimensional, interference, completely redundant, partially redundant) and by task type (identification and search). A target was unidimensional if it differed from others in terms of only one stimulus dimension while all other attributes were

held constant. Multidimensional targets differed from other targets in more than one attribute. A target had interference if irrelevant color was added to the display which interfered with the subject's ability to locate the achromatic target. Color was used as a redundant code if when added to the target it provided more information than was minimally necessary for it to be identified. If knowing the color of a target was necessary but not sufficient to identify it, color acted as a partially redundant code. Identification tasks were measured in terms of accuracy; search tasks in terms of time.

Table 2.2, Range of Percent Difference Scores, summarizes the results of Christ's analyses based on the normalized data, the above definitions of the use of color and the type of tasks the research findings to date had investigated. The data within the table show the range of gain(+) or loss (-) that has been reported with the use of colors as target codes relative to the indicated achromatic coding dimensions. These gains and losses are expressed in terms of the per cent change (Equations 1 or 2) relative to the achromatic codes shown in the first column. The range is based on n number of comparisons.

Christ offers this table only as a starting point to designers considering integrating the attribute of color

Table 2.2

RANGE of PERCENT DIFFERENCE SCORES
for the
USE of COLOR

	Identification Task			Search Task		
	Min	Max	n	Min	Max	n
Unidimensional						
Brightness	+ 29	+ 32	2	+43	+43	1
Size	− 6	+111	6	+40	+40	1
Geometric Shape	− 38	+ 33	11	+ 6	+42	5
Other Shapes	0	+118	6	+30	+63	2
Letters	− 29	− 15	6	+10	+ 7	2
Digits	− 48	+ 26	17	− 3	+42	4
Multidimensional						
Size	− 10	+176	7			0
Geometric Shape	− 28	+202	15	+50	+53	3
Other Shapes	− 2	+ 62	12	+41	+69	6
Letters	+ 4	+ 46	4			0
Digits	− 51	+ 19	6			0
Interference						
Size	− 29	0	14			0
Geometric Shape	− 42	+ 1	4	− 8	− 8	1
Other Shapes	− 43	− 17	4	−10	− 3	2
Digits	− 14	+ 2	7			0
Complete Redundancy						
Size	+ 22	+ 60	3	+33	+32	1
Brightness	+104	+104	1	+32	+32	1
Geometric Shape			0	+21	+32	2
Letters			0	+53	+63	2
Digits	+ 2	+ 2	1	+60	+74	3
Partial Redundancy						
Digits			0	−23	+73	20
Maps	+ 1	+ 1	1			0
Static-Ground Photo	+ 29	+ 29	1	+32	+47	1
Static-Aerial Photo	+ 2	+ 2	1	+17	+17	1
Dynamic-Aerial Photo	+ 3	+ 3	1			0
Dynamic-Aerial TV	+ 3	+ 3	1	− 3	− 3	1

(Christ, 1975, p 563)

into their information display. These results are based solely on works that satisfied Christ et al.'s criteria for their study. Their purpose centered around the recommendation of the use of color in aircraft displays. Identification and search performance were their tasks of interest.

Christ cautions the user on the interpretation of the magnitude of the given ranges. These gain and loss figures were calculated using data reported in the literature. Very few studies mentioned or controlled for the discriminability of the code. Christ suggests this factor may account for the large scores (+ or −) and/or large spread of scores. For example, if the achromatic code was more discriminable than the comparative color code, the loss score may be inflated. However in another study in the same category, if the color code was made more discriminable, the gain score may be inflated. The overall result would then be a large range of per cent difference scores. Few studies have addressed this discriminability of the code or other factors such as ambient lighting, luminance conditions or target background, all of which may have a significant affect on the magnitude of the scores calculated. Christ also pointed out over half of the ranges were derived from two

or fewer studies. Also all studies incorporated in this table used inexperienced subjects. However, when keeping these points in mind, the user can consider this table as a concise summary of what the literature has reported when examining color coding in visual displays.

Christ's secondary objective in searching the literature was to identify gaps in knowledge which would prevent a sound decision on whether to integrate color in aircraft displays. He ascertained that

> the data presented in this report point to many general and absolute gaps in knowledge. There is simply no data available for making some color code-achromatic code comparisons. Furthermore, even when several comparative values are available in the literature, they often have been obtained under relatively restricted conditions. Thus, there is a general need for more parametric research designed to provide data that would allow a broader and a more representative comparison between color codes and various achromatic codes. (Christ, 1975, p. 563)

Christ lists seven specific areas that the research to date has fallen short.

1. The bulk of the data available is derived from subjects inexperienced with the given task. This can restrict generalizations of conclusions.

2. Without exception the subjects were in a laboratory environment where they directed their full attention to the single simple task. Real world displays

usually have to compete for the operator's time with the surrounding environment.

3. The data to date have confounded the display density and exposure time. Both of these variables are of important consideration for display design.

4. Little data are available on how the memory for colors deteriorates during a retention interval.

5. The issue of when a subject can most effectively use a selection criterion based on color of a briefly exposed display has not been fully explored.

6. Some data have suggested color can interfere with the accuracy of identifying and the speed of locating achromatic target attributes. Hardly any data are available to indicate why.

7. Though redundancy of color has not generally been found to aid in identification, does this hold true for practiced operators and/or for targets whose other features are degraded?

Based on these delineations of the voids in experimental research, Christ and his associates designed and performed a series of experiments in attempts to narrow the gaps. The impetus behind their research was the question of whether or not to integrate color into aircraft displays.

Christ and Corso included experience of subject in the experimental design. The eight highly practiced (over one year of training) paid male subjects performed three simple discrete tasks in isolation. The tasks included choice reaction time which required no memory, search and locate, and information-memory task. Codes (letters, digits, shapes and color dots) and density (low and high) were the independent variables with reaction time and accuracy as the dependent variables. The colors used were purple, blue, green, yellow, orange and red.

After analyzing the data collected from 384 trials/subject/condition Christ and Corso's findings were concurrent with earlier results using inexperienced subjects. For the choice reaction (identify) task, digits provided for the best performance while color was consistently associated with poorest performance. In the search and locate task, color was found to be relatively advantageous for target codes in high density displays. Digits, shapes and color did not differ significantly from each other as the best of the four codes to use in an information-memory task that required limited viewing time and short term memory. Their overall conclusion was that experience did not change the effects of using color as a code in simple isolated tasks (Christ and Corso, 1983;

Christ, Teichner and Corso, 1977).

Teichner, Christ and Corso, 1977, felt more realistic laboratory representations of an operational setting needed to be included in the research. Their problem statement was, within a more complex operation will color coding reveal effects not seen in an isolated simple task setting. Using the same three tasks and the same highly practiced subjects as described above in Christ and Corso, 1975, they designed three experiments which combined and integrated the tasks to achieve a more realistic setting. The first experiment combined all three tasks whose displays were all on a panel in front of the subject. A task was initiated at random for the subject to complete. The overall results indicated there were no significant differences between the codes across all the tasks except one. As found in their earlier study, color was found to be the best code for targets in a high density display. The order for codes from slowest to fastest was letter, digit, shape and color dots, with the only statistically significant difference being between letter and color. Again the use of color appeared to be task dependent.

The second experiment was the same as the first with the addition of a same-difference comparison task. The same-difference task involved having the subject determine

whether the display content remained the same as the previous display or had changed (different). New subjects were extensively trained on the isolated tasks. Analysis of that data showed for the original three tasks, color was associated with the slowest response time in the low density identification-memory task. Otherwise the codes were comparable to one another. In the same-difference comparison task colors were responded to with less accuracy and speed than any of the other achromatic codes. Christ, Teichner and Corso attribute the inconsistencies between these findings and those from the first experiment to the differences in practice of the subjects.

A third experiment attempted to put the subject in a more realistic environment. He assumed the role of an air traffic controller (ATC) who had, as a part of his job, the completion of the three original discrete tasks while monitoring and controlling aircraft altitude, speed and heading. The eight original subjects were used. The results showed that with practice to proficiency on the ATC task, the use of color did not affect either job. There was no difference in performance on either the ATC or the simple discrete tasks that could be attributable to color.

The overall conclusion from these experiments as well

as the earlier ones done by Christ and Corso was that color as a code "affects performance in a manner no different from any of the achromatic codes used" (Teichner, Christ and Corso, 1977, p. 33). If there were differences, practice appeared to attenuate those differences. Color does not have a unique quality for coding in a display over other codes in a search and locate type task. It was sometimes associated with the best performance while other times with the worst when using response time and accuracy as the measures of performance (Christ, 1983; Christ, Teichner and Corso, 1977).

Conclusions from the Literature

Almost all studies on the use of color in visual displays have been conducted using search, count, locate and/or identify type tasks. Memory was included in only three: Dooley and Harkins, 1970; Christ, 1975; and Teichner et al., 1977. Most were done in an isolated laboratory setting in which the tasks were simple and discrete. Except for Teichner et al., all generalizations on the use of color have been based on subjects lacking in experience in the task required. Most studies have used letters, digits and geometric shapes as the achromatic codes or brightness with which color was compared for

effectiveness. Without exception the measure of performance of operator was speed and/or accuracy.

Given these conditions, the overall finding was that the relative effectiveness of color as a visual display code depended upon other display conditions, the task and the measure of performance. Generally when performing the simple discrete task of search, locate or identify in an isolated setting, digit coding was superior for the identification task while color was superior for the search or locate task. When these tasks were performed in combination, color appeared equal or superior as a code in reaction and location time and for recall, but inferior in accuracy and speed in same-difference comparison tasks. When code was employed in a combined task which was a subset of the overall job and the operators were proficient at all tasks, color ranked equally with the other three achromatic codes. These conclusions are in conflict with the subjective reasoning that color is obviously better. One might justify the nonconcurrence of the theoretical and experimental findings in how color is presently used as an information code. It may not be taking advantage of the dual processing capabilities of the brain.

Gaps in Knowledge

Color Displays in DSS

There continues to remain glaring gaps in the literature on the knowledge of the use of color in visual displays. The most prominent area was the tasks on which the findings were generalized: search, locate and identify. These are only a small subset of uses of visual displays especially with the proliferation of the affordable computers. Yet the literature continues to dwell only in this narrow field. With computer graphics penetrating almost every sector of the technological world, visual displays are being used everywhere for many different purposes that involve information processing. One of the fastest growing sectors of users are managers and decision makers who are expanding their DSS (Berk et al., 1982; McLamb, 1982). These consumers have the option of having their DSS support color or achromatic coding of their information whether it be graphics or text. Presently, to a limited degree, they also have the choice of a CRT or hardcopy display device which supports the same options. Will color on their visual displays aid managers in their primary function of making decisions at different levels? The subjective literature purports it will. They advocate color will speed up the decision

making process significantly as well as allow for better management recall or memory. For most decision makers time and accuracy are two of their most coveted resources (Kelley, 1979; McLamb, 1982; Miastkowski and Wrege, 1981; Takeuchi and Schmidt, 1980; Watson and Driver, 1983; Zmud, 1979).

The experimental literature has not rigorously addressed the effects of the use of color displays for the decision making process. However if existing results could be generalized, they are in conflict with the popular subjective opinion. The experimental literature concludes the effects of color in visual displays are a function of the task with the overall conclusion that there is not evidence to support color is a better code than other achromatic codes. The question remains systematically unexamined, thus unanswered, on whether color visual displays will aid decision makers in their information processing more than achromatic visual displays.

Researchers in the discipline of Information Systems (IS) agree with those in Human Factors that research is very scarce in the area of the use of visual displays in the decision making process (Benbasat and Taylor, 1978; Dickson et al., 1977; Hanson, 1979; Ives, Hamilton and

Davis, 1980; Mason and Mitroff, 1973; Watson and Driver, 1983; Zmud, 1979). They feel that the DSS have been proliferating with high praise for the technology advancements themselves. However it appears the growth has been in an ad hoc manner with efforts falling quite short of realizing the maximum effectiveness and efficiency of the system design. Mason and Mitroff, noted experts in the discipline of IS research, delineated guidelines for this much needed research. Since a DSS is composed of a person of a certain psychological type, a class of problems, an organizational context, a mode of display and data of a certain form, these are the key variables. In order to design effective displays, these variables must be systematically examined in all possible combinations for main and interaction effects. The majority of research to date has centered around the type of decision maker who uses objective hard facts from a computer printout in a proven formula or model to solve a deterministic class of problems of a nonplanning nature (Dickson et al., 1977; Mason and Mitroff, 1973; Lucas, 1981). Not all DSS are described by these specific characteristics and it is naive to assume that DSS requirements do not vary as a function of the above variables.

There is a growing interest in the mode of visual display of the DSS as it is presumed to have a major impact on if and how the decision maker uses the DSS (Dickson et al., 1977; Lucas, 1981). With the continuing technological advancements, decision makers have choices. They can use a CRT or hardcopy man/machine interface. Their data can be presented on these interfaces in text, graphics or mixed format. This information can be coded in color or achromatic pattern.

Color Graphic Displays in DSS

There is a paucity of research in the area of computer graphics in IS and there is a current need for it especially in a laboratory environment where variables can be controlled and sufficient number of subjects tested (DeMars, 1975; Dickson et al., 1977; Haber and Wilkinson, 1982; Lucas, 1981; Tullis, 1981; Wallace, 1981; Waller, et al., 1982; Watson and Driver, 1983). One reason for this is that prior to 1973 the effects on the human operator were neglected in IS research (Benbasat and Taylor, 1978; Parsons, 1970). The Minnesota Experiments were a start. These experiments addressed the human/machine interface and format options but did not include color. Follow-up and succeeding studies have examined these variables but only two considered color. These were Dooley and Harkins

(1970) and Tullis (1981) discussed earlier. Dooley and Harkins used colored graphics but not as the recognized source of the subjects' task at hand. Tullis (1981) used colored graphics in a training and diagnostic task of a worker which did not involve memory. A CRT was the mode of visual display. Color was compared to achromatic shape coding in the graphical displays.

There is no doubt that color is a very attractive feature and is used extensively in DSS. However there still remains much to be learned regarding when and how to use it to increase a decision maker's performance. Quantitative information about the relative effectiveness on decision maker's output when using color versus achromatic codes in graphics formats is needed so that the benefits color adds to a DSS can be understood and incorporated for its more efficient utilization (Bork, 1980; Chard, 1979; Christ, 1977; Christ and Corso, 1983; Denbrook, 1982; Durrett et al., 1982; Hallworth and Brebner, 1980; Hanson, 1979; Loceff and Loceff, 1981; Taylor, 1979; Truckenbroad, 1981; Tullis, 1981; Waller et al., 1982).

III. METHODOLOGY

Introduction

Research into the understanding of an DSS must consider all parts of the system. However, not all combinations of all parts of the system need be investigated simultaneously or in one study. A referee to Lucas (1981) commented on the general findings of the literature:

> I am afraid that greater understanding in this area (DSS) will be accomplished through many small studies rather than by some major breakthrough (p. 758).

With this in mind a study was designed whose objective was to gain insight on how the attribute of color, when used functionally in computer generated mixed graphics displays, affects the performance of a decision maker who has an engineering background. In this chapter the research method is identified first. Using Mason and Mitroff's definition, the DSS's key parts are then defined with those that are considered variables for this study expounded upon in detail. Pretesting and its results are then described. The chapter concludes with the procedures used in carrying out the study.

Design of Study

Research Method

Van Horn identified four methods of empirical research in the IS area, one of which is a laboratory study using man-machine experiments (Dickson et al., 1977). This method explicitly focuses

> on factors involving the interface between the system and the human decision maker to develop a more meaningful understanding of how persons interact with machine based systems (Dickson et al., 1977, p. 915).

The Minnesota Experiments (Dickson et al., 1977) were all done using the laboratory experiment approach as did Dickson et al. (1977), Lucas (1980) and Zmud (1978) among others. The laboratory setting allows for controls, with benefits that could offset the cost of the simulated environment. Since it was determined the objectives of this study could be satisfied within the framework of a laboratory setting, an experiment was designed using this approach. The particular components of the experiment parallel the key factors of Mason and Mitroff's (1973) DSS definition and are explained next.

Experiment

A MIS is composed a person of a certain psychological type who works a class of problems using data of a certain form displayed on a particular interface within an

organizational setting (Mason and Mitroff, 1973). The DSS of interest in this research involved decision makers who have an academic engineering background. The problem was classified as structured with decisions made under certainty as the data were deterministic in nature. The data were presented in a mixed graphics form either in one of four color alternatives or monochrome. The man-machine interface was computer generated hardcopy. The organizational context was that of management control of resources.

Population. Personal and background differences have been found to influence the way in which decision makers process information and solve problems (Dickson et al, 1977, Lucas, 1980, Mason and Mitroff, 1973, Zmud, 1979). The population of interest is decision makers who have an academic engineering background. It is assumed those students enrolled in the Engineering College, ASU and taking summer coursework as a group do not differ significantly in their decision processes from those in the engineering workforce making the managerial type decisions based on deterministic facts. In fact, many of the students included in the sample have been or were presently a member of the engineering workforce.

The sample population came from those enrolled in

engineering courses taught during the Summer 1983 academic sessions. The experimenter visited six courses (two undergraduate and one graduate class in probability and statistics as well as two undergraduate engineering graphics classes and one graduate decision analysis class) soliciting volunteers. The experiment was briefly explained and a sign-up sheet passed among the students. All but one class instructor offered the incentive of homework points for participation. The majority of each class volunteered. Over 99% of those who signed up for the experiment came on the scheduled day and 120 of them completed the experiment.

The declared academic majors of those who volunteered included aeronautical, biological, chemical, civil, industrial and mechanical engineering as well as physics, mathematics and computer science. The characteristics of interest of the sample used in this study are given in detail in Table 3.1, Characteristics of Engineering Student Sample along with how many of the sample reported those characteristics. These data were extracted from the Background and Experience questionnaire (Appendix A) each subject filled out prior to participating in the actual experiment. The breakdown into levels of each characteristic was based on either a natural split or

Table 3.1

CHARACTERISTICS

of

ENGINEERING STUDENT SAMPLE

Characteristic	Sample Size
Age (years)	
17-22	41
23-28	52
29+	27
Sex	
Female	33
Male	87
College Experience (years)	
<3	37
≥3 and <5	30
≥5	53
Education Level	
<2 years	25
≥2 yrs and not Graduate status	66
Graduate	29
CRT Experience	
Not a User (Never Use)	33
Moderate User (0-6 hrs/wk)	44
Frequent User (>6 hrs/wk)	43
Attitude toward Computer Use	
Uncomfortable	31
Comfortable	48
Very Comfortable	41
Video Game Experience	
Do Not or Seldom Play (0-3 hrs/wk)	39
Moderate Player (>3 hrs/wk)	81

requirements of the analysis techniques used. If a natural split was not obvious, the levels were decided based on a likely natural split as well as the division of the sample characteristics when tabulating them in accordance with the five treatment levels of color. Two criteria were followed. The levels defined were likely. Second, there had to be at minimum two subjects/treatment/characteristic level. This was data dependent because the only independant variable under experimental control was the treatment levels. The other independent variables were based on the sample, who, under the assumption, share the same characteristics as a group as the engineering workforce. The levels of each characteristic factor are defined below.

Age was divided into three levels. A six year span was represented in each of the first two levels 17 to 22 years and 23 to 28 years. The oldest group, 29+ years, included the remaining of the sample, making the categories mutually exclusive and exhaustive.

Sex levels were based on a natural split of female and male. The male/female unbalance was expected as it reflected the proportion of each sex in the engineering student population. This male/female ratio may favor women more than in the overall engineering workforce

because more women are now entering the engineering discipline to train for the workforce.

College experience represented the time each has been exposed to a college level academic train of thought process without controlling for the specific education level. Three levels were defined: less than 3 years, 3 years but less than 5 years, and at least 5 years.

Education level defined the extent of educational experience rather than years in school. Those who have less than two years of college have generally taken only introductory courses. It is assumed those who have had at least two years and were undergraduates have been exposed to a higher level of academic thought process. Graduate students, it is assumed, are taught at a different level that is more demanding and/or creative than undergraduates.

CRT experience followed a natural breakdown. There were those who do not use it (Never Use), interface with it on a moderate basis (0 to 6 hours/week), or frequently use it (more than 6 hours/week).

Attitude towards interfacing with the computer also followed a natural breakout. Subjects reported they were either very uncomfortable or uncomfortable (Uncomfortable), hesitant or comfortable (Comfortable), or

very comfortable (Very Comfortable).

The last variable of interest was video game experience. Many video games require decision making of some type and use the CRT as the human-machine interface. Most commercial as well as the majority of home video games are played in color. To investigate if video game experience affected the performance across the different color treatments, decision makers were categorized in one of two levels. Either a subject did not or seldom played video games (0 to 3 hours/week) or was a moderate player (>3 hours/week). The frequent player category was not included as only two of the sample reported playing more that 6 hours/week. Originally it was intended to break out users as to type played: color or monochrome. However, only two of the users reported playing in monochrome which was not a sufficient sample size for this analysis.

The one characteristic that was considered but not detailed in the table is color blindness. Since one of the key variables of the study is color/monochrome and the treatments were assigned at random, the experimenter excluded anyone who had a potential of being color blind. To test this potential the Dvorine Pseudo-Isochromatic Test was administered to each volunteer. Linksz (1964, p.

238) feels "as a screening device the Dvorine Test is by far the superior one (over the most popular Ishihara Test)." The critical score was 3 or more missed which indicates a potential color vision problem.

Instrument. To satisfy the objective of the study a problem type was chosen that represented a general class of problem and thought process typical of a decision maker with an engineering background. Based on these criteria, the experience of the literature, the contents of several engineering textbooks and the level of the sampled population, an inventory control case study was selected as the content of the decision process.

The logistics management game has been the tool used over the years by MIS researchers to measure the impact of an information system on user performance (Dickson et al., 1977; Lucas, 1980). The game, played at Stanford University and other research institutions for years, involves the subject assuming the role of a decision maker in the management of product logistics. The decision maker, who is given such facts as beginning inventory, demand level, costs, revenues, etc., is asked to maximize profits. A form of this game was used in the Minnesota Experiments. Dickson et al. (1977) also used a modification of this game as the research vehicle in their

nine experiments aimed at gaining insight into the effect of cognitive style and interface output. Lucas (1980) too employed a form of this logistics management game to simulate a DSS when he investigated the impact of the mode of information presentation (hardcopy, CRT, tabular or graphics form). In his conclusions he suggested the game was too complex and recommended it be simplified when used as an experimental tool in further DSS research.

The inventory control problem in its simplest form can be found in many basic engineering textbooks, including those on production control, operations research, decision analysis among others. Because of this and its historic use in MIS research, this problem type was a prime candidate for this investigation. An example problem was found in Hillier and Lieberman (1974) under inventory theory that, when revised, fit the objectives of this research, was similar in nature to the logistic game that had been validated in the literature for DSS research, and was a basic engineering problem.

The selected inventory case study was deterministic instead of probabilistic like those used in literature. Uncertainty would have made the problem more complex as well as limited the population of users to those who were familiar with probability and statistical theory. Lucas

(1980) recommended the problem be simplified. Also, deterministic problems have exact answers which eliminated any bias of the experimenter's own decision making style when scoring the results of the case study. Another reason for making the problem deterministic was the comment:

> One of the most widely observed phenomena of management decision making is the general suppression of uncertainty. In the vast majority of instances, it would appear that uncertainty is not dealt with in any explicit fashion (Morris, 1977, p. 30).

Therefore by making the inventory problem deterministic it became simpler and workable by a larger population, yet without a great loss of reality since most problems are handled deterministically anyway.

The case study is presented in its entirety in Appendix B. Though the problem type is common in many engineering texts, formal academic training on inventory theory was not a prerequisite for volunteers. The mathematical skills and scientific approach underlying most basic engineering coursework were all the tools necessary to work the case study. Calculators, rulers and complex formulas were not needed or used in the experiment. The subjects brought only a pen or pencil.

The case study was divided into four scenarios, each building on the original facts. The arrangement and type

of decisions to be made (questions) as well as the timing are discussed under the measure of performance section. An explanation of the particular format of the case study's data is given in the next section.

Mode of Presentation. The format or mode of presentation in which data are given to the decision maker has elicited considerable attention of researchers as it is a key variable in the DSS's definition. Previous studies have investigated aspects of modes such as narrative versus tabular versus graphical, line versus bar versus pie graphs, among others. This study's basic mode is computer generated mixed graphics. IBM defines mixed graphics as information that is displayed in text and picture format (Benson, 1982). This particular form was chosen for several reasons. Primarily the objective of the study was to investigate the impact of the attribute of color on the performance of the decision maker. Initially, after researching the software and hardware available that supported color, it was discovered that most restricted the use of color to the graphical part of the display. Therefore, those decision makers using computer generated information that had the attribute of color were most likely using graphical output. The assumption is made that even with the emerging available

and affordable technology, the graphics will remain the predominant form of output possessing the attribute of color.

Another reason mixed graphics output was selected as the format was because it is the most commonly used format with the class of problem under study. This is supported by the literature (Dickson et al., 1977; Lucas, 1980) as well as by the software available that is written for this class of problem (Benson, 1982; Kirby, 1979; McLamb, 1982). The most popular of the graphs are bar/column, line and pie charts. This study used the column (vertical) charts because they satisfied the objectives of the study, fit the data presented, are commonly used and most people understand them.

The text part of the display complemented the chart's information. Its purpose was to define the legend of the column charts as well as give detailed information not readily available at a glance on the charts. Each graphic display was designed to stand alone without further documentation.

The graphics standards as specified by Francis (1962), Mclamb (1982) and Schmidt and Schmidt (1979) were followed when designing the mixed graphic displays. The guidelines concerning the text, labels, scale, baseline,

gridlines, graph shape, axis, origin, dimension, spacing and grouping were incorporated into the layout. The patterns and alphanumeric text were computer software generated by the DI (Device Independent) 3000 language, thus satisfying engineering graphics standards.

There were four consecutive scenarios to the case study. The first three defined the facts of the entire case study, with each succeeding scenario building on the previous one. The first three scenarios were essentially alike. They simulated a scenario of a decision maker who had to order the inventory for the summer (Scenario 1). Then the decision maker was informed the fall order had to be placed to the suppliers with the summer order (Scenario 2) and discovered what was ordered in the summer had a significant impact on what was ordered for fall if maximizing total profits was the objective (Scenario 3). In each part, necessary background to the scenario was presented in text form. This information was translated into the mixed graphics displays with one display per inventory item. These graphics were followed by the decisions that needed to be made which were in multiple choice question format. Because the problem was deterministic in nature there could be and was a best objective answer for each question that maximized profit

despite the decision maker's style. In the first scenario all questions were answerable using the mixed graphics displays in that part. For the second scenario's decisions, for which the first scenario's displays were no longer available for review, both the first and/or second mixed graphic information were needed. The third section of the case study had the mixed graphic displays that reflected the changes in the scenario, but whose questions were based on all three parts. The fourth scenario (post test) section of the case study did not contain any new information or mixed graphic displays. It consisted of decisions (multiple choice questions) that simulated a conversation with the next level of management on the conclusions the decision maker made during the analysis portion of the case study.

The primary objective of this research was to investigate the impact of color when used functionally on the performance of the decision maker. Color is used functionally if it is the primary information code (Dooley and Harkins, 1970). The display, shown in Figure 3.1, Mixed Graphic Display, consisted of the legend of graphics and text and the column charts of breakdown (left chart) and total inventory (right chart) dollar figures. There was one display per inventory item/scenario. The color or

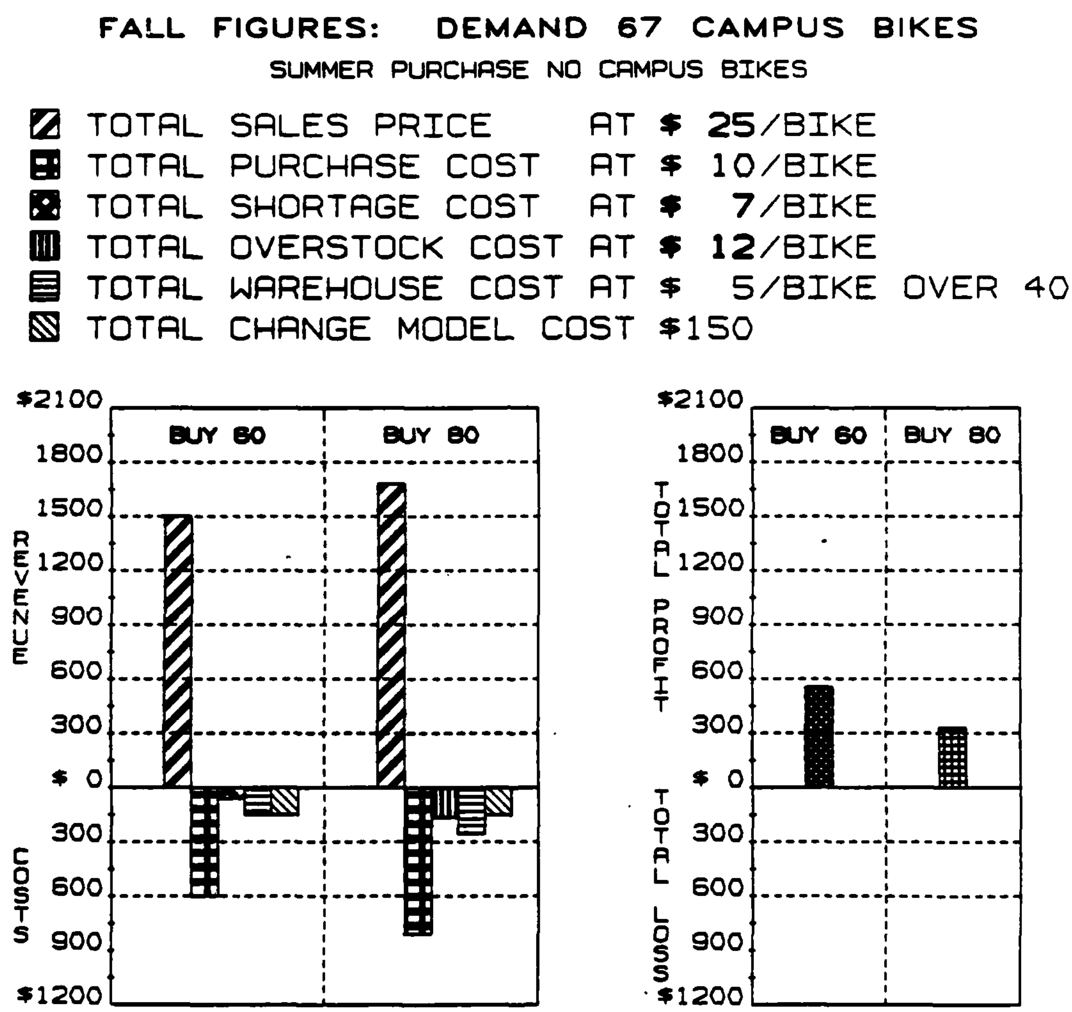


Figure 3.1 Mixed Graphics Display

monochrome code in the legend allowed the decision maker to discriminate between the information available in the breakdown (left) chart.

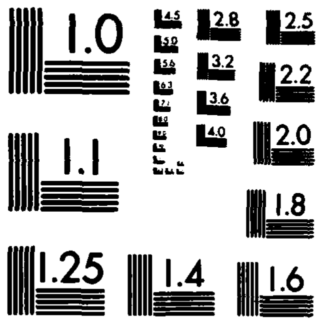
The literature recognizes that there is a void of research in the area of determining what colors should be used. Most studies chose their colors arbitrarily unless they involved situations where specific colors were used because of stereotyping. Using one set of colors and a monochrome configuration, the performance of each was compared. This experiment compared several sets of colors with the monochrome patterns in an attempt to determine if it was the presence of color or a specific set of colors that had an affect, if any, on performance. Four sets of color combinations and one of monochrome pattern were selected as the levels of treatment. Three criteria were followed in determining which color coded what information. First, stereotyping was adhered to where appropriate. Green, if used, always represented a revenue while red coded a cost. Second, two colors of the same hue were not placed adjacent to each other in order to avoid confusion. Finally, the remaining colors were placed at random. Color usage was compatible across all displays as a particular color always symbolized the same information (Keister, 1981).

The first set of colors, subsequently coded CRT in the tables and text, were the primary colors preset on the majority of CRTs that support color. This set of colors is currently the most often used. These included green, red, yellow, blue, dark blue, cyan, black and white. The second treatment, subsequently labeled RND, consisted of a random sample of colors that are available, either primary or user specified, on the CRT colorwheel. The saturation was the same as in treatment one (CRT). This set included green, hot pink, white, turquoise, purple, orange, yellow and black. The third set, subsequently noted as OUT in the tables and text, were consistent with the first treatment: green, red, blue, yellow, white, black, turquoise and cyan. However, the legend squares and column bars were only outlined in these colors, not solidly filled. The fourth treatment, noted as PAS, used a random selection of pastel colors. The intensity of the color was significantly less that that in the CRT or RND treatment. This set was made up of lime green, pink, yellow, white, purple, grey, blue and flesh. The exact identification of each color is discussed in the next section when explaining the interface used.

The fifth treatment was monochrome intensity pattern. Computer software graphics language packages often offer

AD-A145 561 EFFECTS OF ALTERNATIVE CHROMATIC MIXED GRAPHICS DISPLAYS IN DECISION SUPPORT SYSTEMS(U) AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH M S MCCULLY MAY 84 2/4

UNCLASSIFIED AFIT/CI/NR-84-46D F/G 9/2 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

interior intensity patterns that can be used when using a monochrome graphics display. Eight patterns of those available in the DI-3000 graphics language used to generate the displays were selected. Non-similarity in pattern was the criterion used when deciding which eight to use. The patterns, shown in Figure 3.1, correspond to 22, 23, 24, 25, 26, 27, 29 and 36 as shown in the DI-3000 Users Manual (1979), p. A-97. Like the colors, these patterns filled in the legend squares and column bars and were consistent throughout all displays.

<u>Interface.</u> Decision makers utilize various interfaces when interacting with their particular MIS. Hardcopy has been the most prevalent with the CRT gaining due to the availability of both affordable and user friendly hardware, software and service. Although in the near future the CRT will probably be on par or surpass hardcopy as the primary interface, it is assumed the CRT will never replace hardcopy altogether.

The original intent of this research was to utilize the CRT as the man-machine interface. However hardware constraints were encountered which prohibited its use. The case study graphics were programmed using the DI-3000 graphics language software package. This software was selected because it was a Fortran based graphics language

that, while being supported on the Arizona State University (ASU) Harris 800 computer system, was compatible with any interface graphics terminal which had access to the Harris. Due to efficiency of the hard wired terminal, the program (Appendix C) was written, debugged and pretested on the Genesco and Imlac graphics terminals. These interfaces support only monochrome, handle a 9600 baud rate and have a 18 inch screen. The program was written in a way that allowed the decision maker to access any graph in a particular scenario as many times as desired. The average time to paint the graph on the screen was 11 seconds +/- 2 seconds, depending on the number of users accessing the Harris system. The program was self-explanatory even to those who were not familiar or comfortable with interfacing with the computer. The program gave screen instructions, displayed the graphical chart requested by hitting the menu number on the keyboard, denied access to any graph not appropriate to the portion of the case study being worked, counted the total number of graphs viewed and clocked the viewing time of each section. It was written so the decision maker would receive preliminary instructions from the experimenter at the onset, then would work alone with the DSS until ready for the post test. The menu of graphics,

instructions and graphics were displayed on the CRT with the updated scenario and decisions (multiple choice questions) printed on hardcopy.

During the process of writing the software program, portions of it were tested using the Tektronics 4027 as the interface. The Tektronics terminal was the interface originally intended for use in the experiment because it was the only terminal available that accessed the Harris and supported color and graphics. During the time the case study was being programmed, the Tektronics 4027 was in a location that mandated access to the Harris via a modam which allowed a 300 baud rate. Portions of the program were tested as they were developed on the Tektronics 4027 and were found to work the same as on the Genesco and Imlac except for the paint time of the charts. Because of the baud rate (300 instead of 9600) and the raster scan process, the paint time of the chart was five minutes as opposed to 11 seconds. Arrangements were made to relocate the Tektronics 4027 in the ASU Computing Services Computer Graphics Site where it could be hardwired to the Harris, thus have the potential of receiving at a 9600 baud rate.

The Tektronics 4027 terminal was hardwired to the Harris at the time of pretesting. The program was brought

into memory. After the initial instructions and four charts were brought to the screen, the Tektronics terminal would no longer support the case study program. The Manager, Graphics Computing, Mr. J. Howard, analyzed the problem over a two month period. He identified a problem in the DI-3000 software, notified the authors of the language and suggested a fix. It was implemented in the updated version of DI-3000 language but this did not solve the problem. The Tektronics hardware was checked and its memory upgraded. Despite these efforts over a three-month period, the Tektronics 4027 would not support the case study program. In Mr. Howard's expert opinion, it was the hardware of the Tektronics 4027 that was not compatible to the relatively new graphics language despite its device independency. The program executed as intended on both the Genesco and Imlac terminals, but neither of these graphic interfaces supported the option of color. Color terminals with drivers written that would allow access to the Harris mainframe were not scheduled on line until late Fall 1983, which did not fit the schedule of this study. Therefore, the other primary interface used by engineers in many of their decision making sessions, either one on one or in a meeting environment, was chosen for this research. This interface is hardcopy.

The hardcopy mixed graphics displays were generated by building a metafile while painting the display on the CRT screen. A program was written that directed the images to be routed to the Versatec Plotter where the hardcopies were produced. The hardcopy graphics displays were exact images of those given on the CRT interface. Two complete sets of the graphic displays were generated. The first set, whose generator program included the intensity patterns, were used as the monochrome treatment. The second set did not have the legend squares or column bars filled, just outlined in the black ink as a plotter supporting color was not available. The four treatments involving color used this second set as their base.

Color was added to the graphics displays by the experimenter. Using "Pentel" felt tip pens, the legend squares and column bars were filled or outlined in the appropriate schematic as described above for the CRT colors (CRT), the random colors (RND) and the outline colors (OUT). "Venus" colored pencils were used to create the pastel (PAS) treatment. Appendix D identifies the specific patterns and colors used for each treatment.

Organization Context. The final part of the definition of a DSS considers the organizational context. Mitroff and Mason (1973) suggested, based on Anthony's

approach, one way to incorporate this variable is to

> discuss the kinds of problems organizations typically "choose" to solve. This way says that organizational structure is a reflection of the levels of problems the organization considers important to solve (p. 483).

Of the three types (strategic planning, management control, operational control) of major problems organizations typically face, this case study placed the decision maker into the management control position. The objective of the decision maker was to assure that the proper inventory was obtained and resources used efficiently and effectively in the accomplishment of the organization's goal of maximizing total profit. This type of problem satisfied Anthony's definition of management control. Therefore, for this research the organizational context was considered a constant and was defined as managerial control.

Performance Measures

The objective of this study was to analyze how the attribute of color, when used functionally in computer generated graphical displays, affects the performance of a decision maker. King and Rodriguez purport "decision performance assessments reflect the quality of the decision making process that is supported by the MIS"

(King and Rodriguez, 1978 p. 45). Quality of decision making was measured in two of the traditional terms for this research: time and quality of decision.

Time. The research literature considers time to be one of the most important of the measures of performance in either the area of color or DSS. Time was defined in three ways for this study: Total Time, No Memory Time and Memory Time. Total Time was the number of seconds a decision maker used when solving the first three parts or interactive parts of the case study. No Memory Time was the number of seconds used when the task involved no memory but only direct reading from one or more graphs. This was the elapsed time for solving the first scenario of the case study. Memory Time was the amount of time taken to solve those parts of the case study which required memory for some of the decisions. Memory Time was measured as the amount of time (seconds) it took the decision maker to complete the second and third scenarios of the case study. Time was not controlled in the post interview (test) or noninteractive scenario of the experiment.

The experimenter employed a fellow graduate student to measure and record the times. A digital watch was used. Four times were taken: start times for the first

two scenarios, and end times for the first and third scenarios. The same start time was recorded for the entire group when the experimenter signaled them to begin the first scenario. At the end of the first scenario, the subjects were instructed, both verbally before they started as well as in a written direction after the last question of the first scenario, to raise their hand. The timer recorded their time when their hand was raised. At that time the experimenter took the first scenario from the subject and handed out the next two parts. The experimenter signaled the timer when the subject began the second scenario. As instructed for the first part, the subject raised their hand when finished with the third scenario which the timer recorded.

Quality of Decision. As discussed in the literature review chapter, quality of decision (accuracy) is another very pertinent measure of performance when investigating the impact of color or assessing the quality of a DSS. This research measured accuracy in many different ways in order to determine if the affect of color was a function of the type of decision under consideration. The types of decisions and the questions used to measure the performance are defined below. Each operational definition cites the questions which compose the score.

They are referenced by a number which indicates the scenario and specific question of that scenario. All questions are listed in the case study package located in Appendix B. The unit of measure for all scores was per cent.

Total Score was the sum of the scores of all the questions in the case study, 1-1 through 1-9, 2-1 through 2-4, 3-1 through 3-7 and 4-1 through 4-11.

Sum Score was the sum of the scores of the interactive (scenarios 1, 2, and 3) questions of the case study. It included 1-1 through 1-9, 2-1 through 2-4 and 3-1 through 3-7.

Scenario 1 Score was the sum of scores achieved in the first scenario, 1-1 through 1-9. No memory was required to work this section of the case study.

Scenario 2 and 3 Score was the sum of scores achieved in the second and third scenarios (2-1 through 2-4 and 3-1 through 3-7). Some of these questions required memory of what was presented or decided during the previous scenario(s).

Post Test Score was the sum of the points earned on the fourth or interview scenario of the case study (4-1 through 4-11). No new information or previous information were presented, none of the graphic displays were

available for review and the session was not timed.

Direct Read 1 Graph Score was the sum of scores of questions requiring only information read from only one graphic display. Neither memory or computation was necessary to make the decision. This score comprised the points from Question 1-2, 1-3, 1-6, 1-9 and 2-4.3.

Direct Read 2 Graphs Score was the sum of scores of questions needing the data that could be directly read or compared from two or more graphic displays. Neither memory or computation was necessary to make the decision. The score was the sum of points from questions 1-1, 1-4, 2-1, 3-1, 3-1 and 3-7.

Extrapolation 1 Graph Score was the sum of scores of questions requiring manipulation of the information read from one graphic display. No memory was necessary. The score was the sum of points from questions 1-5, 2-3, 3-4 and 3-5.

Extrapolation 2 Graphs Score was the sum of points from questions that needed the computation of the information taken from two or more graphic displays. No memory was required. This score was the sum of the points from questions 1-5, 1-7 and 3-6.

Short Term Memory Score was the sum of points from questions in the interactive scenarios (parts 1, 2, 3)

that required memory of previously presented information and/or decisions. Computation was not necessary. This score was the sum of points from questions 2-2, 2-4.1, 2-4.2, 2-4.4, 2-4.5 and 3-3.

Long Term Memory Score was the sum of the scores of the post test questions that required only memory of information and decisions made during the first three (interactive) scenarios. Computation of the information was not necessary. This score was the sum of the post test questions 4-1, 4-2, 4-7, 4-8, 4-9 and 4-10.

Long Term Memory Extrapolation Score was the sum of scores of the post test questions which needed manipulation of information presented during the interactive scenarios. This score was the sum of points from the post test questions 4-3, 4-4, 4-5, 4-6 and 4-11.

Each question had a correct optimal answer since the problem was deterministic in nature with the overall objective of maximizing profit. This correct answer, which was always one of the available options, was worth five points. Some of the multiple choice decisions had available options that were close to the correct answer. These answers were worth three points because even though the decision was not optimal, it was close. Significantly wrong answers were worth zero points. All questions were

weighted equally in the different scores.

Pretesting

Introduction

The experiment's procedures and instrument were pretested to confirm the timing of the experimental session, the robustness of the computer, the clarity of instructions, the understanding of the graphical legend and charts and the interpretation of the case study questions and answers.

Ten volunteers from a summer graduate engineering decision analysis course partook in the pretesting experiment. Each of the ten sessions was conducted as if it was the actual experimental session with the exception of three factors: the human-machine interface, the environmental setting and the post experiment interview.

Procedures

Each subject was asked to read and sign the volunteer consent form which was written in accordance with the guidelines set by the Council on Human Experimentation. The consent form (Appendix E) delineated the name of the experimenter, the reason behind the research, the purpose and any risks involved with participating in the study, the notification of withdrawal without penalty, and the voluntary consent statement and signature block.

The subjects were then requested to fill out the Background and Experience questionnaire. After each had completed it, they were instructed by the experimenter to read the instructions and the first scenario of the case study which were provided in hardcopy.

They indicated to the experimenter when they were finished reading and ready to begin the interactive decision making part. The experimenter then verbally outlined the instructions again, defined the terms revenue, cost, overage, shortage, profit and loss, and emphasized the decisions were based solely on the information provided in the case study's graphic displays and text. The experimenter also reminded the subject that once a scenario was completed its information and decisions were not available for review, answers could not be changed and the decisions in a particular scenario could be answered in any order. After the experimenter answered any questions, the subject was told to begin.

Since pretesting was accomplished during the period when the CRT was still the projected mode of display, it became the human-machine interface. Both the monochrome Genesco and Imlac graphic terminals were available and supported the program and therefore used. These were located at the ASU Computing Services Computer Graphics

Site. This location was also planned for the actual experiment.

When the experimenter told the subject to begin, a key was struck on the keyboard which initiated the internal computer clock. The clock function was programmed into the case study at specific points to collect and print out the times needed for the analysis. By the way the clock was initiated and stopped, it measured the elapsed time the same for all subjects.

The subject worked the first scenario by referring to the graphic displays that were accessible through a menu selection on the CRT screen. Repeated viewing of any graphics display pertaining to the scenario was an option. A built-in counter was programmed in the software that determined and printed the number of graphics displays viewed during the scenario. If the CRT was the interface, the number of graphics displays viewed would have been a performance variable of interest.

When the subject selected the "Finished with this analysis" menu option, the internal clock marked time and instructions to go on to the next scenario were given. The subject read the new case study facts from hardcopy, viewed the updated graphic displays of his/her choice via menu selection and made decisions that were in the form of

multiple choice and true/false questions given in hardcopy along with the case study text. This process was repeated for the third scenario. When the subject was finished and selected the "Finished with this analysis" option, the internal clock marked time and the program printed the requested times and total number of graphic displays reviewed. The subject was instructed to request from the experimenter the "regional manager's interview" or post test.

This post test was in hardcopy and in the same format as previous decisions. Since the post test was based on information given during the first three scenarios and not on new facts, it did not have graphical information displays for reference. Therefore the subject was asked to move to the free desk next to the CRT to complete the post test. The experimenter stressed to the subject that accuracy of the decisions and not speed was important. The post test was not timed. When the subject was finished the experimenter asked him/her to stay for a post experiment interview.

Post Interview

While the subjects were completing the post test, the experimenter scored the other parts. After they were finished with the post test, the experimenter interviewed

them personally, one on one. The interview consisted of asking them if they encountered any difficulty when interfacing with the computer via the CRT screen and key board. Did they have any problems in understanding any of the written or verbal instructions or scenario text? They were also requested to interpret a graphic display's legend and charts which were presented in front of them in hardcopy. The experimenter chose several questions from those they answered and asked them to explain what the questions meant to them and how they arrived at the answers. These questions included all they had answered incorrectly as well as those with which others had a problem. Any other comments were welcomed. The subjects were thanked for their time and efforts and asked not to discuss the experiment with anyone.

<u>Results</u>

Pretesting was concerned with timing, easiness of working with the CRT and keyboard, clarity of instructions and text, interpretation of graphical displays and questions and difficulty of the questions. The data and comments from the 10 subjects were reviewed which resulted in several changes in the proposed instrument.

Timing of the experimental session was targeted at approximately 45 minutes. Less time may not have allowed

the subject to be absorbed into the case study thus giving it total concentration. Longer time may have resulted in lost interest toward the end of the experiment by the subject, yielding no or invalid data. Pretesting times indicated an average of 55 minutes. It was decided to delete four questions that not only took time but were also found difficult to interpret by several of the subjects.

All subjects, whose attitudes toward interfacing with the computer ranged from very uncomfortable to very comfortable, reported the menu-driven program very easy to control. They commented the instructions were clear and the operation simple. Several said they were distracted by others entering, conversing and leaving the room. To control for this, the terminal used for the experiment was planned to be put in the far corner where it was quieter and the subject's back could face the door.

There did not appear to be any difficulty reading and interpreting the graphical displays. Several of the subjects did not realize at first that all of the questions were based on the information given in the text and on the graphical displays. However, most came to that realization about half way through the first scenario. In efforts to eliminate this problem, the hardcopy

instructions on this point were rewritten for clarity and the experimenter verbalized this instruction twice in the presentation during the actual experiment.

Six questions were consistently missed among the subjects. From their comments, four of these questions were deleted because of the confusion factor. The other two questions were retained. They were answered incorrectly not because of misinterpretation but because of level of difficulty. They were among the most challenging of questions. In fact, all who missed them, when asked how they interpreted them and arrived at their answers, realized their mistake without any input from the experimenter.

The difficulty of the questions were analyzed by scoring the 10 sets of case studies and figuring the average of each scenario by itself and in total. The four deleted questions were not included. The averages calculated ranged from 80% to 91%. This was higher than the population targeted average of 75%. However the higher scores were expected because of the pretesting sample used. They were all graduate students currently enrolled in a graduate decision analysis course that was leading up to covering a similar problem type. Those subjects were getting into that mode of thought prior to

being exposed to the experiment. Too, all of them partook in the experiment within one hour after their two hour decision analysis class. Based on this analysis it was determined the questions were at the proper difficulty level that would yield data analysis that would be meaningful to the research question.

During the pretesting of the procedures and instrument, the Tektronics 4027 color terminal became available for hardwiring into the Harris 800 mainframe. It was then the problems were encountered and researched as previously discussed under Design of Study Interface. When it was decided to use hardcopy as the human-machine interface, several changes had to be made to the instrument. The text instructions were rewritten to reflect the use of hardcopy instead of the CRT interface. These changes were minor. Pretesting was accomplished by recontacting three of the subjects and asking their interpretation of the instructions. No difficulty was encountered. Too, hardcopy graphic displays had to be computer generated with the color treatment added where appropriate. The hardcopy graphic displays did not have to be pretested as they had already been used in the original pretesting session for interpretation analysis.

After all the changes were incorporated into the

design, it was determined to proceed with the laboratory experiment. Analysis of the design and the pretesting information gave confidence that the experiment would yield data which, when analyzed properly, would allow the researcher to gain insight on the research question.

Experimental Procedures

Introduction

The same procedures were followed when conducting the experiment as were suggested during the pretesting phase with several additions. The subjects were selected and scheduled for the experiment, the treatments were completely randomized in their assignment to the subjects, the subjects were tested for color deficiency, and the experiment run.

Sample Selection

The experimenter visited six of the engineering prefixed courses that were taught during the summer sessions, 1983 at ASU. The experiment was defined as a case study that involved making decisions based on information provided on graphic displays. It was emphasized that the only prerequisites were that they be enrolled in the engineering course and would volunteer an hour and a half of their time and effort. No special skills or knowledge was necessary. It was open to all and

incentive homework points were offered by five of the instructors. The treatments (monochrome and color) were not mentioned. A signup sheet was passed around for volunteers to schedule a time convenient for them. In all of the classes, over 90% volunteered.

Scheduled Environment

The experiment was administered at eight different times, four times each in the morning and afternoon. Three different rooms were used as the environment, each used at least once in the morning and in the afternoon. One of the rooms was a classroom, the second a lecture hall classroom and the third a graphics laboratory classroom. All rooms had flat tabletop desks. The classroom and lecture hall desks were 29 inches high. The chairs in these two rooms were identical and were the contoured backed chairs with seat height 17.25 inches from the floor. The laboratory desks were drafting tables, which when the top was in the horizontal position as it was during experimentation, it was 37 inches high. A backless stool was used whose seat was 26 inches from the floor. The lighting was 440 LUX (morning) and 500 LUX (afternoon) in the classroom, and a constant 430 and 460 LUX respectively in the windowless lecture hall and graphics laboratory. The noise level in each setting was

only that made by the participants themselves and the experimenter. A quiet study environment was maintained. Signs posted on the rooms' doors were successful in deterring outside distractions. The only major distractions in the rooms were the movements of the experimenter and those of the subjects when leaving after they had finished.

Randomized Treatments

Two techniques were used in attempts to randomize the treatments among the sample, yet at the same time trying to achieve a balanced design and control for exogenous variables. The signup system allowed the experimenter to estimate how many to expect at each session. To insure an ample and an equal number of treatments were available at each experimental session, three or more was added to the signup number to make it a multiple of five. It was divided by five to determine how many of each treatment were needed. Using a random number table, the treatments were ordered and assigned a sequential number. These treatments were placed blank side up in sequential order on every other desk in the experiment room, starting from the speakers right to left, front to back. The rooms were vacant when the treatments were distributed. The subjects arrived and sat at any vacant desk that had a case study

on it. The experimenter insured none of the subjects touched the instruments before being told to do so. A post experimental survey of a random sample from each session indicated the subjects were unaware of the different treatments both prior to and after completing the experiment. Based on these two procedures, it is assumed that every subject had a equal chance of receiving any one of the treatments, which according to Kendall and Stewart (1971) satisfies the principle of randomization of treatments.

This process insured that at each of the eight sessions approximately the same number of each treatment was administered. This allowed the experimenter to control for exogenous variables that may have otherwise biased the inferences gained from the data. The key exogenous variables of concern were competition, verbal instructions and subject questions, courses from which the subjects came, desk type, noise, lighting, time of day, day of week and month.

A competitive atmosphere often exists in a classroom in a testing environment. This had a potential of affecting the performance of a worker in different ways. No matter what the nature of the competitive affect, it is assumed it did not favor any particular treatment since

each type were represented in approximately equal quantities.

The written instructions were controlled but the verbal ones may have varied each time they were given by the experimenter even though the same written outline was used. Even when the exact words were verbally read, inflections of certain instructions had a potential of affecting the outcome of the experiment. Too, after the verbal directions were given, the experimenter entertained all questions which in themselves could have influenced the outcome. Again since all treatments were administered in approximately equal numbers in each session, it is assumed this factor did not bias any one treatment.

Most sessions were scheduled to be convenient for the volunteers, which usually meant right before or after their classes. Therefore the majority of subjects/session came from the same class. To wash out any bias that may have arisen from this homogeneity of the sample, the five treatments were administered in equal numbers. The objectives of this research did not include generalizing on a population of engineers discriminated by specific coursework background.

Like the particular course experience, the desk type, noise, calendar date, day and time of day were not

considered variables of interest to this research. However, logistics did not allow these factors to be considered or assumed constants. Therefore, by insuring a balance of treatments under each condition, it is assumed that these factors, though varied, showed no favor to any one or more of the five treatments.

Color Deficiency

As the subjects selected a seat for the experiment, the experimenter screened each for potential color deficiency. If a subject missed three or more of the patterns shown in the plates of the Dvorine Psuedo-Isochromatic Test, they were asked to forego participating in the experiment, given class incentive points and requested to leave. Nine males and zero females, which is approximately proportional to the general population (DeMars, 1975) were asked to leave. There was never more than two in a session who did not identify the required number plates.

Experimental Procedures

After all subjects were seated, had been tested for potential color deficiency, and completed the Volunteer and Background and Experience forms, the formal experiment began. Instructions were given; the first scenario worked, timed and collected; the next two scenarios given,

completed, timed and collected; and finally the post test taken.

Time was allotted for each subject to read the instructions and first scenario of the case study. After all had done this, the experimenter verbalized the instructions, defined the terms revenue, cost, overage, shortage, profit and loss, and emphasized the decisions were based solely on the information provided in the case study. All questions were then entertained. After the last was answered, the experimenter again reminded the subjects that the decisions were based only on the information provided; the scenarios built on each other but were not available for review once completed, and decisions could be made in any order but once marked could not be changed. She also stressed that as soon as they finished the first scenario they were to raise their hand because timing was important to the study. The group was then directed to turn over the case study packet located in the upper left of their desk and begin. The recorder marked the start time.

As the subjects worked, the experimenter and timer both scanned the room to insure that each were using the graphics in their activities. Since this direction did not appear clear during the pretesting phase, the

experimenter wanted to insure it was for the experiment. During all the sessions, it was not apparent to either the experimenter or the timer that any of the subjects misunderstood this direction.

The timer continually scanned the room for those who finished. His objective was to mark the clock time when the subject raised their hand indicating the scenario was done. The timer monitored for time as well as any indications that a subject was finished but did not follow the directions concerning raising their hand. Three subjects fell into this category, which the timer caught and recorded the clock time after each looked up and before they started gazing and wondering what to do next. With only about 15 subjects in each session, the timer located himself so he could see all the subjects papers without moving about the room. It is assumed since the sessions were small, the same timer and his digital watch used, and the same procedures followed, the timing was consistent over all treatments and sessions. If it was not, the inconsistencies were random over the treatments.

When a subject raised their hand indicating the end of the first scenario, the timer recorded the time and the experimenter went to the front of the subject's desk. She took the first scenario's text and graphics displays, and

reviewed the pertinent instructions for the next two scenarios in a quiet one-on-one conversation. The second scenario was handed face down to the subject and the third part laid in the upper left corner of their desk. When the subject turned the second scenario over to work it, the experimenter signaled to the timer to record the start time.

The subject was instructed to work through the second scenario and when finished, lay it (text, decisions and graphics) face down on the empty desk adjacent to them. They were to continue right on to the third scenario that was located at the top left corner of their desk. From the observations made by both the experimenter and timer, no one had difficulty following these instructions.

As for the first scenario, each subject was directed in writing and verbally to raise their hand when finished with the third scenario. The timer followed the same procedures in determining the time to record and marked it as he did for the first scenario. For this time, every subject appeared to follow the requested directions concerning raising their hand.

When the subject finished the interactive part of the case study, the experimenter approached the front of the desk. She cleared the desk of everything and handed the

subject the post test. The subject was verbally told the post test was based on the previous information and decisions. This part was not timed and accuracy or quality of decision was the important concern. Each were thanked for their generous participation and were asked to leave quietly after they finished the decisions and turned over the paper on their desk. They were reminded not to discuss the experiment with anyone.

The experimenter randomly selected three from each session to interview after the experiment. She did this to discover if any had knowledge of the key treatments being tested either before of after the experimental session. Twenty-three were unaware their graphics displays were different than others in the room. One noticed on his way out that another had colored graphics displays while his were black and white.

After each session the experimenter checked through each turned-in case study. This was done only to determine if all the questions in each part were answered, not for scoring. Only one problem was encountered. One subject only completed the first two questions of the post test. Since this did not occur during the last session, the data were considered invalid, were not used and the subject was not counted against the total needed for that

treatment type. From a quick scan of the case studies from the last session, which brought the total to 120, the data collection, thus the experimental phase of this research, was finished.

IV. ANALYSIS AND RESULTS

Introduction

The objective of this research was to gain insight on how the attribute of color, when used functionally in computer generated mixed graphics displays, affects the performance of a decision maker who has an engineering background. A laboratory experiment was created that simulated a decision making environment in an attempt to gather data that, when analyzed properly, would lend insight to the problem. This chapter focuses on the analysis portion of the research. Its development follows the chronological order of the types of analysis done.

The first section discusses the analysis of the problem statement which resulted in defining hypotheses which, when tested statistically, would satisfy the objective of this research. This section delineates all the formal hypotheses considered in this project. The second section presents the analysis techniques others have employed when engaged in similar types of research. Since the analytic tools selected depended not only on the hypothesis, but also on the properties of the data used to test the hypothesis, the criteria underlying the tools are then examined in the third section in light of both the hypotheses and data. Based on the degree of satisfaction

of the criteria, the appropriate analytic tools are identified and applied to the data. The results are given and conclusions offered for all hypotheses under consideration in the fourth section. The chapter concludes with the results and conclusions derived from the overall trends in the data.

Hypotheses

The problem statement was broken down into many parts and rewritten in a form that was conducive to rigorous objective analysis. Each part, or hypothesis, narrowed the general problem. This allowed an in-depth study on different aspects of the problem as well as an overall analysis of the problem when considering all the results as a whole. This multi-hypotheses approach gives strength to the conclusions derived from the results.

The performance of the decision maker with an engineering background is what was measured during the experiment. Performance was defined in terms of time to make decisions and quality or accuracy of decisions. Three different times were defined: total time to accomplish the interactive part of the case study, time to complete the first scenario which did not require any memory, and time to work those graphics supported scenarios that mandated memory to make many of the

decisions.

Quality of decision was the other quantitative measure of performance. It was subdivided into 12 types, each of which was operationally defined in the Performance-Accuracy section of the previous chapter and are defined again in the Performance Measure Selection section of this chapter. They included total score of the case study, sum score of the interactive parts, no memory interactive score, memory interactive score, post test score, direct read from one graph score, direct read from more than one graph score, extrapolation from one graph score, extrapolation from more than one graph score, short term memory score, long term memory score and long term memory extrapolation score.

Hypotheses were formulated that incorporated each of the performance variables with every combination of treatment and characteristic factor. These are listed in Table 4.1, Research Formal Hypotheses. Each hypothesis is two sided with the equality of performance in the null (Ho) and the inequality in the alternate (H1) hypothesis. Subsequently only the null hypotheses are stated, with the alternatives implied. The general mathematical form is

$$H(k)o: \quad u(i) = u(j) \quad \forall ij, \quad \begin{array}{l} i=CRT, RND, OUT, PAS, B/W \\ j=CRT, RND, OUT, PAS, B/W \\ k=\text{Characteristic Factor} \end{array}$$

H(k)1: u(i) $\neq$ u(j) for at least one i,j
k Constant

where

u = average performance of those using a particular mixed graphics display treatment. Performance was measured in 15 ways, 3 of which were a function of time, 12 a function of decision quality.

i,j = the particular graphics treatment. The five treatments under study included

- CRT Standard CRT color solid filled graphics
- RND Random selection of available CRT color solid filled graphics
- OUT Standard CRT color outlined graphics
- PAS Pastel color solid filled graphics
- B/W Random selection of available monochrome intensity pattern graphics

k = characteristic factor whose effect on the performance was accounted prior to testing the hypothesis. There were seven characteristic factors, each subdivided into several levels. These factors included age, sex, college experience, education level, CRT experience, attitude toward interfacing with a computer and video game experience.

This generic hypothesis was therefore tested for 15 alternative measures of performance and 7 characteristics factors. This represented 105 different tests. Table 4.1 specifically defines the null hypothesis for each of the forms.

Literature's Analytic Approach

The two basic analytic tools employed by those reporting their experimental studies on the effects of color in the literature were a form of the Analysis of Variance (ANOVA) and the t-test (Christ et al., 1977;

Table 4.1

RESEARCH FORMAL HYPOTHESES

H(k)o: u(i) = u(j) $\forall$i,j i=CRT,RND,OUT,PAS,B/W
j=CRT,RND,OUT,PAS,B/W
k=Characteristic Factor
H(k)1: u(i) $\neq$ u(j) for at least one i,j k Constant

TIME

Total Time (TT)
H(k)o: u(TT of CRT)=u(TT of RND)=u(TT of OUT)=u(TT of PAS)
=u(TT of B/W)

No Memory Time (NT)
H(k)o: u(NT of CRT)=u(NT of RND)=u(NT of OUT)=u(NT of PAS)
=u(NT of B/W)

Memory Time (MT)
H(k)o: u(MT of CRT)=u(MT of RND)=u(MT of OUT)=u(MT of PAS)
=u(MT of B/W)

QUALITY OF DECISION

Total Score (TS)
H(k)o: u(TS of CRT)=u(TS of RND)=u(TS of OUT)=u(TS of PAS)
=u(TS of B/W)

Interactive Score (IS)
H(k)o: u(IS of CRT)=u(IS of RND)=u(IS of OUT)=u(IS of PAS)
=u(IS of B/W)

No Memory Interactive Score (NS)
H(k)o: u(NS of CRT)=u(NS of RND)=u(NS of OUT)=u(NS of PAS)
=u(NS of B/W)

Memory Interactive Score (MS)
H(k)o: u(MS of CRT)=u(MS of RND)=u(MS of OUT)=u(MS of PAS)
=u(MS of B/W)

Post Test Score (PS)
H(k)o: u(PS of CRT)=u(PS of RND)=u(PS of OUT)=u(PS of PAS)
=u(PS of B/W)

Direct Read 1 Graph Score (D1)
H(k)o: u(D1 of CRT)=u(D1 of RND)=u(D1 of OUT)=u(D1 of PAS)
=u(D1 of B/W)

Table 4.1 (Cont.)

RESEARCH FORMAL HYPOTHESES

Direct Read 2 Graphs Score (D2)
H(k)o: u(D2 of CRT)=u(D2 of RND)=u(D2 of OUT)=u(D2 of PAS)
=u(D2 of B/W)

Extrapolation 1 Graph Score (E1)
H(k)o: u(E1 of CRT)=u(E1 of RND)=u(E1 of OUT)=u(E1 of PAS)
=u(E1 of B/W)

Extrapolation 2 Graphs Score (E2)
H(k)o: u(E2 of CRT)=u(E2 of RND)=u(E2 of OUT)=u(E2 of PAS)
=u(E2 of B/W)

Short Term Memory Score (SM)
H(k)o: u(SM of CRT)=u(SM of RND)=u(SM of OUT)=u(SM of PAS)
=u(SM of B/W)

Long Term Memory Score (LM)
H(k)o: u(LM of CRT)=u(LM of RND)=u(LM of OUT)=u(LM of PAS)
=u(LM of B/W)

Long Term Memory Extrapolation Score (LE)
H(k)o: u(LE of CRT)=u(LE of RND)=u(LE of OUT)=u(LE of PAS)
=u(LE of B/W)

Hitt, 1961; Hoisman et al., 1963; Kanarick and Petersen, 1971; Munns, 1968; Schutz,1961; Smith, 1963; Smith and Thomas, 1964; Smith et al., 1965; Saenz and Riche, 1974; Tullis, 1981; Wedell and Alden, 1973). The underlying assumptions which, when satisfied, give these analytic tools their validity, were very seldom addressed in any of the literature findings. The ANOVA or t-test was applied to the data, the statistics generated and tested for significance at the .05 significance level, with conclusions and generalizations offered. For most of the studies, the choice of analytic tool was based solely on the hypothesis under scrutiny without consideration of the properties of the data.

In the literature when both time and quality of decision were of concern, the most common procedure was to treat these measures of performance as separate variables without detailing justification. However in Hitt's (1961) study of the relative effectiveness of using selected abstract coding in a map reading task, he combined them, using decision accuracy/time as the measure of performance. Hitt reported two main reasons for doing this. First, the product moment correlation between accuracy and speed was -.85. Based on this, he concluded the two measures were highly linearly correlated. The

second reason Hitt gave was that this transformation was needed to achieve a normal distribution of the data and homogeneity of variance. When each measure was plotted alone, each formed a highly skewed distribution. This transformation resulted in an interpretable variable whose data satisfied the conditions underlying the ANOVA technique, which Hitt used.

Hoisman et al. (1963) investigated the effects of color coding on subject performance when completing six categories of map reading tasks. They also examined the relationship between time and accuracy. However their correlation analysis of their pilot as well as actual studies' data suggested no significant correlation. Based on this, they assumed the two performance variables independent, thus analyzed each separately. The remaining of the reports examined time and/or accuracy separately without inferring a justification.

Congruent with the research and mathematical literature, the analytical techniques used in the study reported here involved the determination of the form of the performance variable, the definition and validation of the mathematical model and the analysis of results.

Analysis

Introduction

Each performance variable, color treatment and characteristic combination's data were analyzed. This was necessary to test the hypotheses stated in Table 4.1. Based on the conclusions derived from the hypotheses testing, inferences could be made on how using the attribute of color in graphics displays affects the performance of a decision maker with an academic engineering background.

This section explains how the data were prepared for input, what software supporting the data manipulation was selected, how the data were analyzed and what results and conclusions were derived from analysis of the above stated hypotheses.

Data Preparation

As detailed in the methodology chapter, the raw data were contained in the case study packages used as the instrument during experimentation. When the experiment was completed, all of the case studies were scored in one sitting by the researcher. This helped control for consistency in scoring across all studies. As noted earlier, the questions were deterministic in nature and the correct response was objective and one of the

available multiple choice options. After all 120 case studies were scored, a random sample of 25 were selected for rescore to confirm consistency. No inconsistencies were cited.

The individual scores in each case study were coded into a software file to be used for the data manipulation. A hardcopy was printed and compared for errors against the original case studies' scores, times and volunteer characteristics. After transcription corrections were made, the researcher was confident the data used in the analyses were accurate and valid.

Statistical Software

Because of the volume of data and the breadth of the analysis, a mathematical software package was sought to aid in the manipulation. Two packages were found that were accessible to the experimenter and supported the needed techniques. These were the Statistical Analysis System (SAS) and the Biomedical Computer Programs P-Series (BMDP). Not only did they have the necessary analytical techniques, but they also had the option of interaction with each other. This allowed the output from one package's technique to be used as input to the other's procedures which at times was needed. The main programs were written in SAS which had the option of calling BMDP

subprograms. Both packages were accessed via interactive keyboard/CRT interface on the IBM WYLBUR mainframe located in the ASU Computing Services.

Analysis of Properties of Data

Introduction. Based on the literature review and the hypotheses under consideration the ANOVA appeared to be the analytic technique that would generate sufficient statistics which would allow inferences to be made concerning the effect color has on the performance of a decision maker. However the mathematical literature warns on using this model without first ascertaining if the data support it (Box, Hunter and Hunter, 1978, Neter and Wasserman, 1973). The statistics the model provides maintain their sufficiency only when the data in the model meet certain conditions. These conditions, often called assumptions, are independence of the error term, normality of the residuals, homoscedasticity and additivity of the factors.

An analytic algorithm was developed for the analysis portion of this research based on the recommendations of the mathematical literature. This approach considered not only the characteristics of the hypotheses under investigation but also the properties of the data used for the analysis. The algorithm is shown in Figure 4.1,

Analysis Strategy. Initially a measure of performance was determined (i.e. Time vs Score) based on correlation analysis of the time and score data. Using this measure as the dependent variable, a specific tentative ANOVA model was formulated. The four underlying criteria of independence, normality, homoscedasticity and additivity were then examined separately to validate the use of the model. Based on the results of the analysis of the assumptions, the appropriate model and statistics were chosen for analysis of the affects of color on performance. Results and conclusions were generated based on the analysis of the statistic on how using the attribute of color in graphics displays affects a decision maker with an academic engineering background.

The data analysis portion of this chapter follows the outline given in Figure 4.1. Under each heading the criterion is defined and examined for the effect on the aptness and alternatives to the model if the data depart from it. Next the data for each research hypothesis are analyzed with all results shown in an accompanying table. Conclusions are offered for each criterion.

<u>Performance Measure Selection.</u> Initially the form of the performance measurement (dependent variable) had to be determined. If evidence existed to reject the null

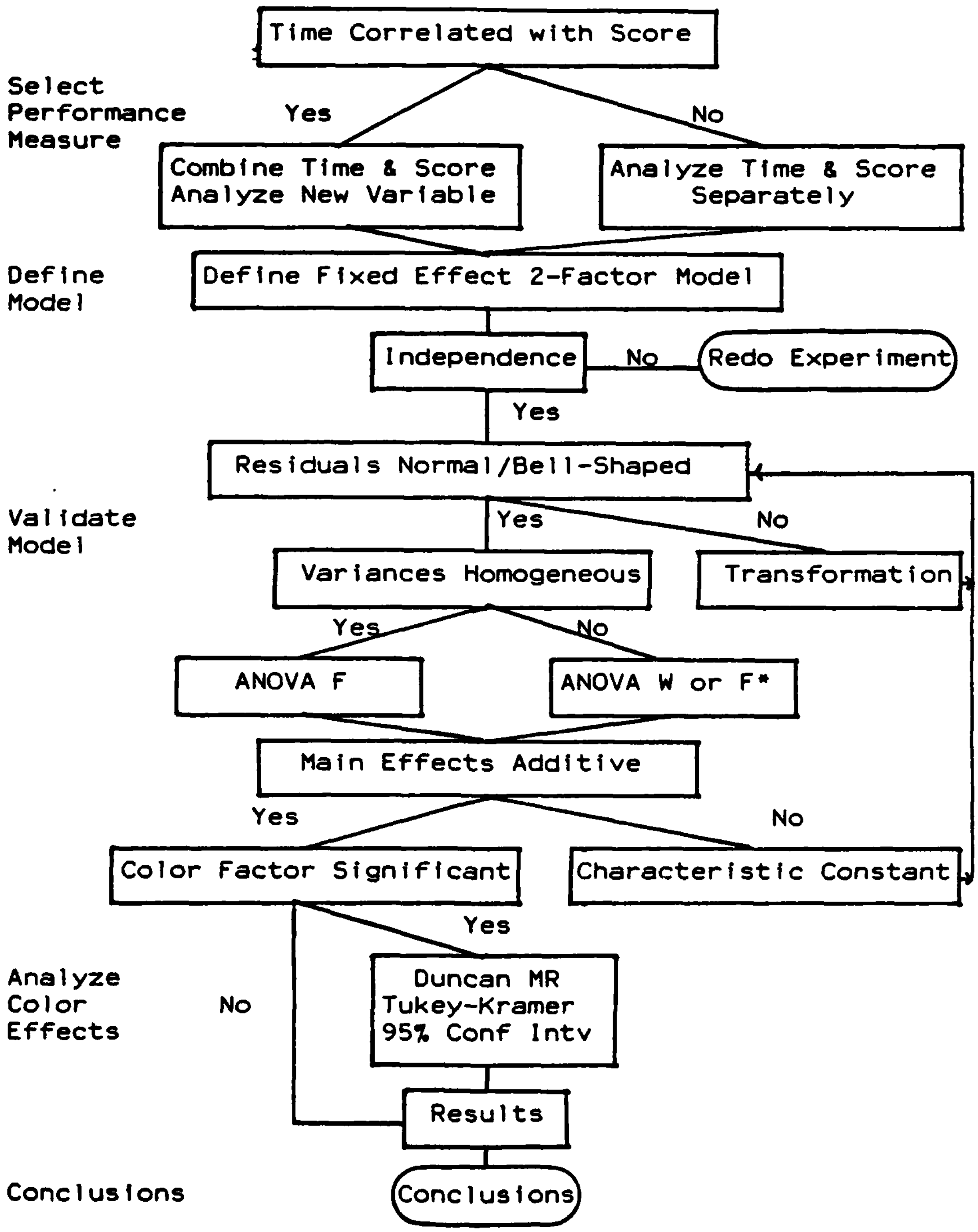


Figure 4.1 ANALYSIS STRATEGY

hypothesis that time and score were uncorrelated, the two were combined as score/time for the remaining analysis. If there was insufficient evidence to reject the hypothesis that the correlation between time and score was zero, then each was considered as a separate measure of performance throughout the remaining analysis.

Correlation analysis was done between the time and score performance variables to determine the form appropriate for the remaining analysis. Since both the time and score performance were measured at the interval level, the Pearson Product Moment Correlation technique was used to generate the correlation coefficient statistic r (Conover, 1971). The hypothesis Ho: $\rho = 0$ was tested. The results are shown in Table 4.2, Correlation Analysis: Score vs Time.

In Table 4.2, the times recorded and available for analysis head the columns. Total time was the overall time in seconds used to complete those parts of the case study where graphics displays and scenario text were accessible. No Memory Time was the time lapsed to finish the first scenario, in which recall was not needed for the decision making. Memory Time was the time passed when completing Scenarios 2 and 3. These two scenarios contained a mix of question types, many requiring memory

Table 4.2

CORRELATIONS ANALYSIS: SCORE vs TIME

Ho: $\rho = 0$

p Values
Reject Hypothesis if $p \leq .05$
Sample Size N=120

TIME / QUALITY	Total Time	No Memory Time	Memory Time
Total	.81	.37	.49
Sum	.87	.07	.16
Scene 1	.46	.52	
Scene 2&3	.40		.41
Post Test	.78	.08	.26
D Read 1	.67		
D Read >1	.81		
Extrap 1	.28		
Extrap >1	.37		
Short Mem	.77		.92
Long Mem	.81	.39	.25
Long Mem Extrap	.60	.10	.50

* Rejected hypothesis at $p \leq .05$

of previous information or decisions. The post test was not timed.

The quality of decision variables (scores) are listed in the first column of rows. Each is operationally defined below when the results of the analysis are discussed.

The entries in the table are the computed levels, often referred to as the p values. The null hypothesis tested was the correlation coefficient is equal to zero or that there is no correlation between the time variable at the top of the column and the score variable at the beginning of the row. If the computed α level or p value was greater than an appropriate significance level, the conclusion was that insufficient evidence existed to reject the hypothesis of zero correlatior. Entries are presented only for those time/score combinations which had relevance.

Total Score of the entire case study did not appear correlated with the overall time it took to complete it. There is an .81 probability of being wrong if the decision of correlation is made. Nor was there evidence to suggest total score was correlated to either no memory time (p=.37) or no memory time (p=.49).

Sum Score, the sum of the scores obtained when the

graphics displays and text were accessible (Scenarios 1, 2 and 3), were not related to time at $\alpha = .05$ in any of its forms. Sum Score and Total Time were measured over the same three scenarios, yet if correlation was accepted, there is an .87 probability of being wrong. There was insufficient evidence at the .05 significance level to support that either Sum Score and No Memory Time (p=.07) or Sum Score and Memory Time (p=.16) were linearly correlated.

Scenario 1 score, which defined quality of decisions requiring direct read, compare or extrapolate and no memory did not show a linear relationship with either Total (p=.46) or No Memory (p=.52) time. This score and No Memory time were measured over the same scenario, yet there was not evidence to support correlation (p=.52).

Scenario 2 and 3 score was comprised of scores from those case study parts. It represented quality of decisions which required direct read, extrapolation, comparison and/or memory. This score did not appear to be systematically related to either the Total Time (p=.40) or Memory Time (p=.41) which covered the same scenarios.

Post Test score was that computed from the post interview when no new information was given and none of the text or graphics displays were accessible. These

decisions required long term memory and/or extrapolation of data. There was no evidence to suggest the time spent using the graphics displays whether as a whole (p=.78) or broken down (p=.08 and p=.26) consistently affected the quality of decisions made during the post interview.

Direct Read 1 graph score was the sum of scores of all decisions requiring reading from only one graph. Memory or computation of data was not required. The questions included were 1-2, 1-3, 1-6, 1-9 and 2-4.3. These were the simplest decisions. The Direct Read score did not appear correlated to the total time (p=.67).

Direct Read >1 graph score was the sum of scores of all decisions requiring reading from more than one graph which did not require memory or manipulation of data. Questions 1-1, 1-4, 2-1, 3-1, 3-2 and 3-7 defined this variable. Based on a p=.81, it was assumed there was not a significant linear correlation between this score and Total Time.

Extrapolation 1 graph score was the composite score of decisions which required the decision makers to extract information from the text and/or one graphics display, and manipulate it to arrive at the solution. Memory was not required. The questions were spread across the first three scenarios to include 1-5, 2-3, 3-4 and 3-5. There

was not evidence to suggest a correlation between this measure of quality of decision and Total Time (p=.28).

Extrapolation >1 graph score was defined the same as the above score except the decision was based on information given in two or more graphics displays. Questions 1-5, 1-7 and 3-6 composed the score. The data did not give support (p=.37) to correlation between this quality score and Total Time.

Short Term Memory score was the sum of points from decisions in the interactive scenarios (Scenario 2, 3 and 3) which required memory of previously presented information and/or decisions. These questions included 2-2, 2-4.1, 2-4.2, 2-4.4, 2-4.5 and 3-3. Computation was not necessary in these decisions. Based on the p values of .77 and .92, it was assumed a linear relationship did not exist between Short Term Memory score and either Total or Memory Time respectively.

Long Term Memory score was a measure of quality of decision made during the post test. It is divided into two categories. The first, Long Term Memory, pooled the points of the post test questions which required only memory and no further calculation or manipulation of data. These questions were 4-1, 4-2, 4-7, 4-8, 4-9 and 4-10. The second category, Long Term Memory Extrapolation,

included those post test decisions which needed further data or information manipulation for solutions. These questions included 4-3, 4-4, 4-5, 4-6 and 4-11. Based on the derived p values, neither of these scores appeared correlated with any of the time categories.

The overall results indicate there was insufficient evidence to conclude that time and quality of decision were linearly correlated. Therefore it is assumed for this research the covariance is zero between the time and any quality score measure of performance. Thus each were analyzed separately.

Define Model. A majority of the treatments and characteristic factors were classified qualitatively. Performance, a function of time or decision score, was a quantitative variable measured at least on the interval scale. The statistical literature suggests the 2-factor fixed effects linear ANOVA approach. It allows comparison of multiple means based on several factors simultaneously without loss of power (Montgomery, 1976) and yields sufficient statistics if all of the underlying assumptions are satisfied (Box et al., 1978). The ANOVA chosen was

$$y(ikl) = u + \tau(i) + \beta(k) + \tau\beta(ik) + \epsilon(ikl)$$

i=Treatment Level
k=Characteristic Level
l=Observation Number

where y(ikl) was the (ikl)th performance (time or decision quality score) observation and u was the overall mean performance effect. The $\tau(i)$ represented the true effect of the ith level of color treatment (B/W, CRT, OUT, RND or PAS). The $\beta(k)$ was the true effect of the kth level of the characteristic factor (age, sex, college experience, education level, CRT attitude and experience and video game experience). $(\tau\beta)(ik)$ represented the effect of the interaction between the ith color treatment and the kth characteristic factor and $\epsilon(ikl)$ was the random error component of the (ikl)th observation. Table 4.3, Variables for Analysis lists the variables, levels and sample sizes for each term defined in the models used for analyses. The treatment factor is balanced across all levels at 24 sample points. This was controlled by the experimenter. The sample size in the levels of all the characteristic factors was a function of the random sample. The performance variables are listed with sample size of 120 points for each of them.

A fixed effects model was directed by the problem

Table 4.3

VARIABLES for ANALYSIS

Treatment Factor (τ_i)	Sample Size
Color	
CRT Color (CRT)	24
Black & White (B/W)	24
Outline (OUT)	24
Random (RND)	24
Pastel (PAS)	24

Characteristic Factor (β_k)	Sample Size
Age (years)	
17-22	41
23-28	52
29+	27
Sex	
Female	33
Male	87
College Experience (years)	
<3	37
$\geq$3 and <5	30
$\geq$5	53
Education Level	
<2 years	25
$\geq$2 yrs and not Graduate status	66
Graduate	29
CRT Experience	
Not a User (0-3 Never Use)	33
Moderate User (0-6 hrs/wk)	44
Frequent User (>6 hrs/wk)	43
Attitude toward Computer Use	
Uncomfortable	31
Comfortable	48
Very Comfortable	41
Video Game Experience	
Do Not or Seldom Play (0-3 hrs/wk)	39
Moderate Player (>3 hrs/wk)	81

Table 4.3 (Cont)

VARIABLES for ANALYSIS

Performance Variables y(ikl)	Sample Size
Quality Score (Percentage Points)	
Total Score	120
Sum Score	120
Scene 1 Score	120
Scene 2/3 Score	120
Post Test Score	120
Direct Read 1 Graph	120
Direct Read >1 Graph	120
Extrapolation 1 Graph	120
Extrapolation >1 Graph	120
Short Memory	120
Long Memory	120
Long Memory Extrapolation	120
Time (Seconds)	
Total Time	120
No Memory Time	120
Memory Time	120

statement. The objective was to determine if the presence of color in a mixed graphics display affected the performance of a decision maker. Each color combination was chosen for a particular reason as described in the earlier chapter. Since each color combination available on a CRT display interface did not have an equal probability of being selected, the model had to be fixed as opposed to randomized. Inferences made from the results are constrained to the color treatments employed.

Model Validation

Introduction. Though the ANOVA model appeared to yield information that would allow insight into the research problem, it had to be examined for its aptness (Box et al., 1978; Neter and Wasserman, 1973). The statistics this model provides maintain their sufficiency only when the data used in the model meet certain conditions. These conditions, often called assumptions, are independence of the error term, normality of residuals, homoscedasticity and additivity of factors. Each of these conditions was examined in light of how departures affect the aptness of the model. Using the analytic tools suggested in the mathematical literature, the data were analyzed to determine the degree of satisfaction of the criteria. The assumptions are

addressed in the order shown in Figure 4.1.

Independence. Independence of the error terms can not be relaxed. The error term is considered independent if the subjects in each of the levels of factor groups are randomly and independently selected so the outcomes of any subject had no effect or influence on any other's outcome within or among any of the groups. Proper randomization of the experiment introduces independence in the assignment of the treatments to subjects. The resulting experimental errors may then be regarded as independent. Usually proper randomization is achieved by randomly assigning the order in which the observations are taken such that every treatment had the same chance of being applied to each eligible subject (Kendall and Stuart, 1966; Montgomery, 1976).

As detailed in the Experimental Procedure section of the methodology chapter, each color treatment was randomly assigned so that every level of treatment had the same probability of being applied to each volunteer subject. Therefore it is assumed the random errors are independent and this criterion satisfied.

Normality of Residuals. The ANOVA is relatively robust against the criterion of residual normality for a fixed effects model. The point estimates of the factor

level means are unbiased as long as the distribution of the residuals form a bell-shaped curve. The ANOVA uses the F test for the equality of factor level means, which is relatively insensitive in terms of power and significance to the lack of normality (Berenson, Levine and Goldstein, 1983; Box et al., 1978; Neter and Wasserman, 1973). If there is reason to believe that the data do not form at least a bell-shaped curve, the literature recommends applying an appropriate transformation of the data to make the distribution normal or bell-shaped.

Several criteria were used to determine if there was evidence to support the rejection of the hypothesis of residual normality. The residuals for each cell of each ANOVA model matrix under consideration were calculated. The Shapiro-Wilk test was applied to test for normality. If the derived α value in each of the cells where the hypothesis was tested was greater than .05, the hypothesis was not rejected, the residuals were assumed normal and the entry in Table 4.4, Normality of Residuals, recorded as N. For those cells in a model for which the derived α was less than .05, the kurtosis and skewness statistics were analyzed for bell-shaped determination. If the skewness was less than 1.0 and kurtosis less than 2.5, the

distribution was assumed normal/bell denoted by N/B in Table 4.4 (Lewis and Ford, 1983). If the skewness was less than 2.0 and there were two or fewer of the 10 or 15 cells that had kurtosis greater than 2.0 but less than 7.0, the distributions of the cell residuals were assumed bell-shaped. These were coded B in the Table 4.4. Those which did not satisfy any of the above were not assumed either normal or bell-shaped but skewed. These were coded S in Table 4.4.

Table 4.4 presents the results of testing the residuals for normality or bell-shaped characteristics. The paired independent factors head the column and the performance measure considered starts the row. Using the criteria stated above, the statistics from the Shapiro-Wilk normality test were analyzed. As shown in the table, there was insufficient evidence to reject the assumption the residuals were distributed normally or at least bell-shaped for all performance measures except one. The exception was the model whose performance measure defined quality of decision derived from direct reading of only one graphics display. Analyzing these cell distributions indicated all were highly skewed to the left. This was anticipated since it was the simplest of decisions and all the scores clustered at the maximum

Table 4.4

NORMALITY of RESIDUALS

Ho: Residual Are Normal or Bell-Shaped

Factors / Performance	Color & Age	Color & Sex	Color & Col Ex	Color & Edu Lv	Color & CRT Ex	Color & Attit	Color & VGame
QUALITY							
Total	B	N	N/B	B	N	N/B	N
Sum	N	N	B	N	B	N	N/B
Scene 1	B	N/B	B	N/B	B	B	B
Scene 2&3	N	N/B	N	N	N	B	B
Post Test	B	N/B	B	N/B	N	B	B
D Read 1	S	S	S	S	S	S	S
D Read >1	B	N/B	N/B	N/B	N/B	B	B
Extrap 1	B	N/B	N	B	N/B	N/B	N/B
Extrap >1	B	N/B	B	B	N/B	B	B
Short Mem	B	B	B	B	N	B	N/B
Long Mem	B	N/B	N/B	B	B	B	N/B
Long Mem Extrap	N/B	N/B	N	B	N/B	N	B
TIME							
Total	B	B	B	B	B	N/B	N/B
No Memory	B	B	N	N/B	N	B	N/B
Memory	N	B	N	B	N	N	N/B

N Normal
N/B Normal/Bell-Shaped
B Bell-Shaped
S Skewed

score. Thus the model using the direct read from one graph as the performance measure was not considered for further analysis. Instead it is used to validate the assumption the subjects interpreted the graphics displays correctly. Therefore, for the remaining analyses, it is assumed the normality of residual assumption is satisfied to the degree necessary for valid use of the 2 factor fixed effect linear model for all hypotheses under scrutiny.

Homogeneity of Variances. Unless the sample sizes are equal, departures from the criteria of homoscedasticity could seriously affect the level and power of the F test (Berenson et al., 1983; Box et al., 1978; Brown and Forsythe, 1974; Montgomery, 1976; Neter and Wasserman, 1974). Unequal variances invalidates the use of the ANOVA model when sample sizes are unequal across treatments.

There are several methods available to determine if the variances are equal. One such model is the Modified Levene's L test which is relatively robust to the normality assumption (Berenson, et al., 1983) and does not require equal sample sizes.

The literature suggests two ways to handle the problem of unequal variances: transformation of data or

recalculation of the F statistic of the ANOVA. A transformation is a viable option if the residual data are not distributed normally as it usually restores normality as well as stabilizes the variances. However it has been found that if a transformation is applied to normal data in attempts to equate the variances, normality is almost always destroyed (Berenson et al., 1983; Neter and Wasserman, 1974). If this occurs, the approximation by Welch or Brown-Forsythe can be employed directly on the original data in lieu of the usual ANOVA to compare the means (Berenson et al., 1983; Brown and Forsythe, 1974).

Welch and James both extended the Welch asymptotic confidence interval approach for the two-means problem to more than two groups. Brown and Forsythe (1974) studied the behavior of the F statistic (ANOVA), W statistic (Welch), the J statistic (James), and F* statistic (Brown-Forsythe modified F). From their results they concluded that when the variances were equal, both the W and F* are close in power to the F. The choice between W and F* depends on whether the extreme means are thought to have extreme variances (W) or not (F*). When the variances were not equal, the F statistic showed marked deviation from its nominal size (3-17%). Therefore they feel

> Considering the small loss in power in using F* or W rather than F even when the population variances are equal, protection against F's distortion in size recommend the use of F* or W. The asymptotic approximations of the critical values of F* and W are valid when each has at least 10 observations and are not unreasonable for W down to 5 observations (Brown and Forsythe, 1974, p. 131).

Levene's test of equal variance was used to test the hypothesis that all variances in a specific model were equal. This test did not require the data to be normal or the sample sizes equal. The results shown for each model are presented in Table 4.5, Homogeneity of Variances.

The matrix is the same as used in Table 4.4. The independent factor combinations head the columns and the performance measures start the rows. The table entries are the computed or p values. The null hypothesis tested was the variances in a specific model's cells were all equal. If the computed α or p value was greater than the significance level, set at .05, the data did not contain sufficient evidence to reject the hypothesis of equal variances.

Those tests where there was evidence to support rejecting the equal variance hypothesis are denoted by an asterisk (*). Since the data were assumed at least bell shaped, failing this underlying assumption meant using the Welch W or Brown-Forsythe F* statistic for further

Table 4.5

HOMOGENEITY of VARIANCE

Ho: Cell Variances Equal

p Values
Reject Hypothesis if $p \leq .05$

Performance \ Factors	Color & Age	Color & Sex	Color & Col Ex	Color & Edu Lv	Color & CRT Ex	Color & Attit	Color & VGame
QUALITY							
Total	.62	.13	.06	.08	.80	.56	.15
Sum	.17	.07	.11	.07	.79	.51	.15
Scene 1	.00*	.00*	.15	.00*	.01*	.15	.03*
Scene 2&3	.33	.78	.31	.75	.94	.35	.21
Post Test	.33	.03*	.06	.00*	.03*	.13	.21
D Read >1	.25	.23	.23	.40	.15	.39	.73
Extrap 1	.38	.53	.52	.68	.84	.15	.62
Extrap >1	.95	.39	.85	.10	.39	.45	.51
Short Mem	.20	.19	.68	.74	.19	.06	.03*
Long Mem	.04*	.06	.01*	.01*	.11	.03*	.00*
Long Mem Extrap	.52	.21	.88	.26	.36	.13	.93
TIME							
Total	.02*	.56	.16	.56	.73	.45	.15
No Memory	.06	.50	.00*	.15	.08	.02*	.20
Memory	.04*	.61	.48	.34	.19	.23	.69

* Rejected hypothesis at $p \leq .05$

analyses. For those models whose data did not give sufficient evidence to reject the equality of variance, further study was done by applying the two-factor fixed effects ANOVA technique using the F statistic.

Table 4.6, ANOVA Statistic for Analysis, presents the particular statistic used based on the equality and size of variance. If the variances were assumed equal as shown in Table 4.5, the ANOVA F statistic (F) was used for further analysis. If there was evidence to suggest hetrogeneity, most cell sample sizes were at least 10 and the extreme means did not have proportional variance, the Brown - Forsythe F* statistic (F*) was employed for further analysis. Finally, if the variances of a model were not considered homogeneous, a majority of the cells sizes numbered less than 10 or the extreme means were estimated to have extreme variances, the Welch W statistic (W) was used. As a result, Table 4.6 displays the ANOVA statistic used to resolve the research questions. Of the 98 hypotheses, 80 were examined using the F statistic, 15 using the Welch W statistic and 3 using the Brown-Forsythe F* statistic.

Additivity of Main Effects. Factor effects are considered additive if there are no significant factor interactions. Each factor can be described separately by

Table 4.6

ANOVA STATISTIC for ANALYSIS

Factors / Performance	Color & Age	Color & Sex	Color & Col Ex	Color & Edu Lv	Color & CRT Ex	Color & Attit	Color & VGame
QUALITY							
Total	F	F	F	F	F	F	F
Sum	F	F	F	F	F	F	F
Scene 1	W	F*	F	W	W	F	W
Scene 2&3	F	F	F	F	F	F	F
Post Test	F	W	F	W	W	F	F
D Read >1	F	F	F	F	F	F	F
Extrap 1	F	F	F	F	F	F	F
Extrap >1	F	F	F	F	F	F	F
Short Mem	F	F	F	F	F	F	F*
Long Mem	W	F	W	W	F	W	F*
Long Mem Extrap	F	F	F	F	F	F	F
TIME							
Total	W	F	F	F	F	F	F
No Memory	F	F	W	F	F	W	F
Memory	W	F	F	F	F	F	F

F ANOVA F Statistic
F* Brown-Forsythe Statistic
W Welch Statistic

analyzing the factor level means of the factor main effects. However if significant interaction exists between the main factors, the main effects statistics have little practical meaning. When this is the circumstance, the literature suggests holding one factor constant while applying the ANOVA on the others in order to draw conclusions on the non-constant main effects (Montgomery, 1976; Neter and Wasserman, 1974).

Additivity was checked in all 98 cases by testing the hypothesis that the interaction effect was zero. If there was insufficient evidence to reject this hypothesis, the main effects were considered additive and analyzed by the suggested statistic shown in Table 4.6.

However if the data indicated the interaction effect was not zero, the characteristic factor was held constant at each of its levels and the entire analytical process repeated starting at the testing of residual normality. The statistics derived from the repeat analysis on the new subset of data were used for analysis of the main effects. Inferences based on this data were constrained to the particular level of the characteristic factor.

Table 4.7, Additivity Analysis, presents the p values which were calculated when testing the hypothesis the interaction factor $(\tau\beta)_{ik}$ was zero or the main effects

Table 4.7

ADDITIVITY of MAIN EFFECTS

Ho: $(\tau\beta)_{ij} = 0$

p Values
Reject Hypothesis if $p \leq .05$

Factors / Performance	Color & Age	Color & Sex	Color & Col Ex	Color & Edu Lv	Color & CRT Ex	Color & Attit	Color & VGame
<u>QUALITY</u>							
Total	.24	.01*	.43	.17	.91	.48	.27
Sum	.37	.02*	.52	.07	.96	.32	.32
Scene 1	.65#	.09#	.40	.03**	.82#	.07	.27#
Scene 2&3	.24	.03*	.72	.23	.91	.63	.46
Post Test	.43	.03**	.34	.34#	.61#	.57	.53
D Read >1	.65	.04*	.65	.21	.92	.87	.82
Extrap 1	.86	.87	.32	.04*	.22	.35	.77
Extrap >1	.44	.27	.89	.49	.71	.51	.06
Short Mem	.06	.24	.46	.46	.93	.36	.29#
Long Mem	.59#	.61	.82#	.39#	.43	.06#	.35#
Long Mem Extrap	.56	.11	.19	.31	.73	.62	.23
<u>TIME</u>							
Total	.42#	.35	.03*	.89	.92	.47	.75
No Memory	.98	.39	.03**	.84	.59	.41#	.77
Memory	.16#	.31	.20	.84	.92	.43	.70

* Rejected Hypothesis based on F Statistic at $p \leq .05$
** Rejected Hypothesis based on inflated F Statistic at $p \leq .05$
\# p Value using (1.25)(ANOVA F)

were additive. The entries in Table 4.7 were calculated in two ways, depending on the method used for analysis: ANOVA F or Welch and Brown-Forsythe (see Table 4.6: ANOVA Statistic for Analysis). If the ANOVA F technique was applied, the F statistic reported for the interaction term was used to determine the tail probability or p value. They are indicated by a plain or * entry in Table 4.7. The * indicates the data gave evidence to reject additivity of the main factors at the .05 significance level. For these cases, the characteristic factor was held constant at the different levels and the entire analysis procedure repeated, beginning at the step of testing for normality of residuals. These cases included Color x Sex vs Total Score, Color x Sex vs Sum Score, Color x Sex vs Scenario 2 & 3 Score, Color x Sex vs Direct Read >1 Graph Score, Color x College Experience vs Total Time, and Color x Education Level vs Extrapolation 1 Graph Score.

For those cases which earlier analyses indicated Welch W or Brown-Forsythe F* should be applied, the entries in Table 4.7 were calculated by inflating the ANOVA F. The available Welch method only tested main effects. Brown and Forsythe (1974) reported their analyses showed the ANOVA F statistic fluctuated as much

as 17% when the small samples had large variances. To account for a possible fluctuation that may affect the testing of the additivity hypothesis, the ANOVA F statistic was inflated by 25% before determining the p value. If the data contained evidence that rejected additivity, the inflated F would be more sensitive in detecting it. These entries are noted in the table by # or ** symbols. The # indicated there was no evidence to reject additivity, therefore the Welch W or Brown-Forsythe F* statistic was used to analyze the main effect of color. For those with **, interaction may exist. For these models, the characteristic factor was held constant at the different levels, and the entire analysis procedure repeated starting at testing the residuals for normality. Three cases fell into this category: Color x Sex vs Post Test Score, Color x College Experience vs No Memory Time, and Color x Education Level vs Scenario 1 Quality Score.

Analysis of Color Effects

Preliminary Color Effects Analysis. After the assumption of additivity of main factors was satisfied, the statistics derived from the ANOVA model became sufficient estimators of the parameters of primary interest. They were analyzed to determine if the average performance of the decision maker with an academic

engineering background was affected differently by one or more of the color treatments. Only the statistics estimating the color treatment parameter (τ i) were of concern to this research. If the ANOVA F, F* or W indicated no significant difference among the treatment means at the .05 significance level, the result was to fail to reject the null hypothesis under consideration. For that particular performance variable for which the variability due to the characteristic factor (βk) was accounted, the results supported the conclusion that there was no difference in performance attributable to the color of graphics display used.

For the three cases where the variances were not equal and additivity could not be assumed (Sex vs Post Test, Col Exp vs No Memory, Edu Lev vs Scene 1), the method of analysis was reevaluated during the repeat of the analysis procedure. These hypotheses contain more than one p value/entry in Table 4.8, Analysis of Main Effects of Color. These p values correspond to those calculated for each level of the characteristic factor under consideration and are reported in the same order as the levels are listed in Table 4.3, Variables for Analysis. That is for Sex, the pairs contain the p values for female and male respectively. The College Experience

Table 4.8

ANALYSIS of MAIN EFFECT of COLOR

Ho: $\tau i = 0$

p Values
Reject Hypothesis if $p \leq .05$

Factors / Performance	Color & Age	Color & Sex	Color & Col Ex	Color & Edu Lv	Color & CRT Ex	Color & Attit	Color & VGame
<u>QUALITY</u>							
Total	.19	.01* .12	.22	.26	.15	.13	.03*
Sum	.22	.02* .07	.12	.38	.15	.17	.05*
Scene 1	.24	.13	.13	.11 .06 .61	.43	.21	.01*
Scene 2&3	.32	.04* .17	.36	.55	.32	.36	.11
Post Test	.19	.04* .20	.15	.04*	.38	.10	.05*
D Read >1	.53	.20 .06	.38	.92	.26	.25	.22
Extrap 1	.99	.72	.74	.24 .05* .26	.45	.47	.41
Extrap >1	.64	.28	.23	.18	.11	.21	.05*
Short Mem	.20	.27	.54	.48	.73	.88	.48
Long Mem	.28	.08	.67	.24	.04*	.22	.39
Long Mem Extrap	.57	.23	.41	.57	.35	.27	.11
<u>TIME</u>							
Total	.25	.84	.79 .03* .50	.71	.55	.32	.70
No Memory	.86	.94	.75 .06 .81	.87	.86	.58	.94
Memory	.39	.24	.15	.39	.40	.12	.44

* Rejected hypothesis at $p \leq .05$

triple reports the p value for Under 3 Years, 3 and less than 5 Years and 5 Years or more. Less than 2 Years, 2 or more Years and Undergraduate Status, and Graduate is the order for the p values for Education Level. These p values were based on the statistics derived during the analytic reevaluation of the revised data set.

When only the post test performance of males was investigated, the data supported using the ANOVA F. However, the variances were not equal among the female groups of color treatments. The sample size in each color treatment mandated the W statistic. When the Color x College Experience vs No Memory Time was analyzed one experience level at a time, analysis indicated Welch W be used for two of its levels: less than 3 years and more than 5 years college experience. The ANOVA F was sufficient for the middle level of at least 3 years but less than 5 years college experience. After examining the data for the three levels of the Education Level x Color vs Scenario 1 Quality Score, unequal variances and sample size suggested Welch W for the less than 2 years level, Brown-Forsythe F* for the at least 2 years and Undergraduate. The data of the graduates satisfied the ANOVA F criteria, thus it was used.

<u>In-depth Analysis.</u> If the ANOVA statistic indicated

a significant difference ($p \lesssim .05$), further in-depth analysis was done in efforts to discriminate among the means. This was accomplished by applying one of three methods: Duncan Multiple Range Test, Tukey-Kramer Test or 95% Confidence Interval Graphical Inspection. If the original model under consideration assumed equality of variance and additivity, the Duncan Multiple Range Test was applied to the data. This test is reported to be one of the most powerful (Box et al., 1978; Montgomery, 1976). If the variances were assumed homogeneous but additivity was rejected and further analysis was done using the data from a particular level of a characteristic factor, the Tukey-Kramer Test was applied. This test is a modification of the Tukey Test, and unlike both the Duncan and the Tukey tests, it is robust against unequal cell size. Since the characteristic factors were not controlled for balance, when additivity was rejected, the cell sizes were guaranteed to be unequal. If the model did not satisfy homoscedasticity, the data were analyzed by building 95% confidence intervals about all the means. These confidence intervals were plotted and a subjective analysis was graphically conducted. The results from Duncan and Tukey-Kramer are also given as estimates of the groupings to support the graphical analysis. Each of the

hypotheses that was rejected at the .05 significance level was analyzed using one of these three methods and are discussed separately.

Prior to presenting the final results, one model was selected to follow through the entire analysis process. The model defined, validated and analyzed using the developed analyses strategy shown in Figure 4.1 and detailed above. The relevant portions of the analysis strategy diagram are filled in with the appropriate statistics and p values derived from the models data. This is shown in Figure 4.2, Analysis of Color x Age vs Scene 1&2 Quality Score. The analysis is discussed next in detail.

The example model examined is Color x Age vs Scene 2&3 quality score. Independence was assumed based on the procedures employed in randomizing the color treatments among the three age groups under investigation.

The normality of residuals was tested by examining the p values derived from applying the Wilk-Shapiro test on the residual data in each cell of the 2-factor fixed effects model. These are shown below in Figure 4.3.

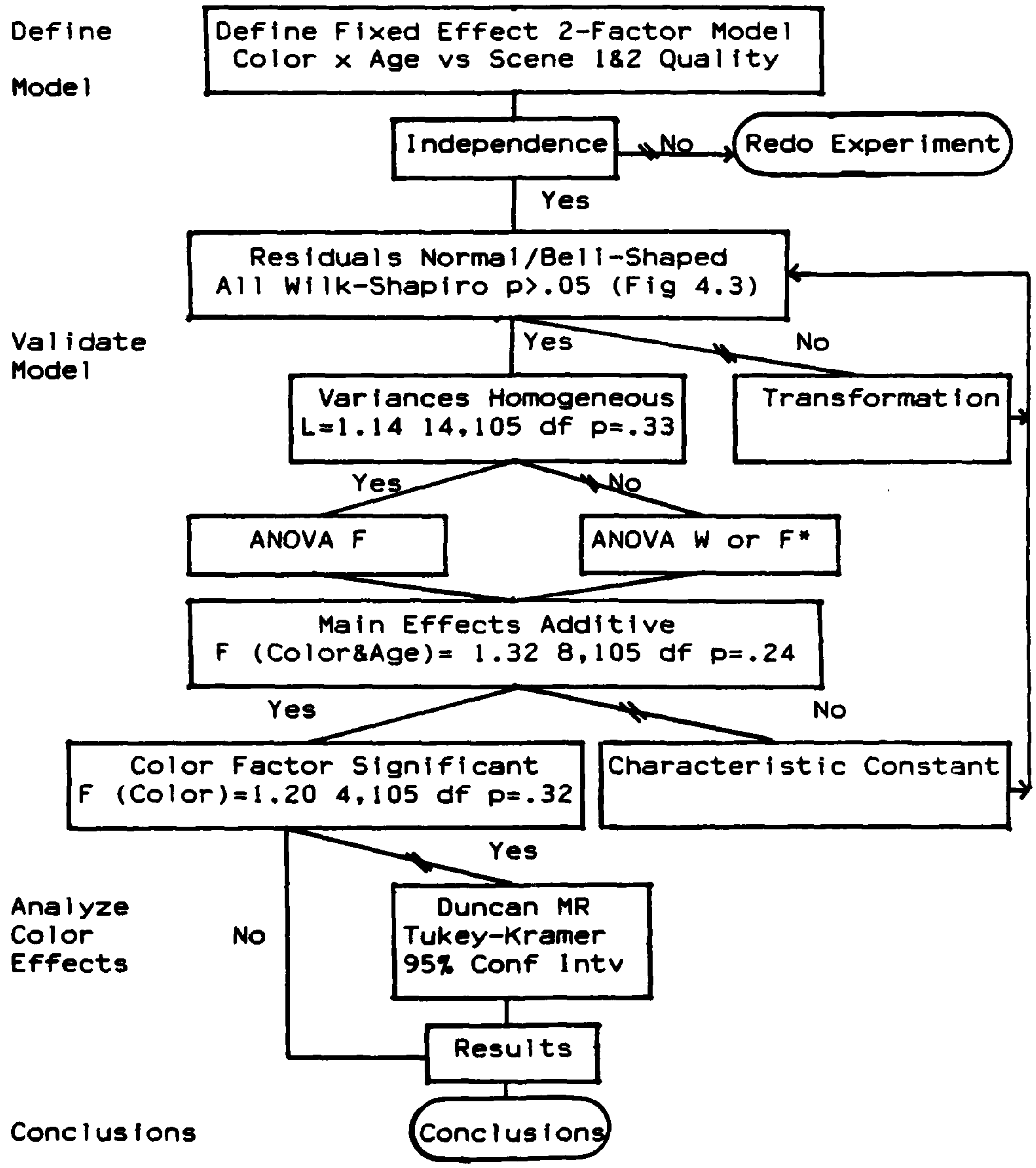


Figure 4.2 ANALYSIS of COLOR & AGE vs SCENE 1&2 QUALITY

Color / AGE	B/W	CRT	OUT	RND	PAS
17-22	.84	.17	.10	.10	.23
23-28	.46	.61	.82	.83	.09
29+	.30	.91	.42	.99	.22

Figure 4.3 p Values Derived from Wilk-Shapiro

As seen in the table, all the p values were greater than the .05 significance level. There was insufficient evidence to reject the hypotheses that the residuals in each cell of the ANOVA model were normally distributed. Therefore a transformation was not needed.

Hetroscedasticity was tested next. The p value resulting from applying the Modified Levene's test was compared to the .05 significance level. Levene's L=1.14 with 14,105 degrees of freedom corresponded to p=.33. The data contained insufficient evidence to reject the hypothesis that the variances were homogeneous across all cells of the defined model. Equal variances implied the derived ANOVA F statistics were sufficient for testing the next criterion of additivity of factors.

Additivity of main factors was tested by examining the F and its associated p value for the interaction term of age and color. The 2-factor ANOVA method was applied via the SAS program which yielded F=1.32 with 8, 105

degrees of freedom. The associated p value was .24. The data contained insufficient evidence to reject the hypothesis the main factors of color and age were additive. Based on this result the ANOVA F was validated as a sufficient statistic for analyses of the hypothesis under scrutiny.

After the variability in performance attributable to age was accounted, did color affect the Scene 1&2 decision quality? This question was answered by analyzing the ANOVA F for the main factor of color derived from applying the ANOVA on the 2-factor fixed effect model. The resulting F was 1.20 with 4, 105 degrees of freedom which corresponded to a p value of .32. Since this p value was greater than .05, the significance level, the conclusion was made the color factor was not significant. If it had been, the data would have been further analyzed using Duncan's Multiple Range as the variances were assumed equal as well as main factors of color and age additive.

This same procedure was repeated for all of the formalized hypothesis listed in Table 4.1. The results of these analyses are delineated next.

Results

Introduction

The overall objective of this research was to examine

how the attribute of color, when used functionally in computer generated mixed graphics displays, affects the decision maker who has an academic engineering background. One hundred eleven models were analyzed to gain insight into the research question. The results of testing the null hypothesis that performance as measured by quality or time of decision was not affected by the color of the mixed graphics display are shown in Table 4.8, Analysis of Main Effects of Color. Each entry in this table is the computed α or p value derived from the appropriate ANOVA statistic as discussed previously.

Color Factor Insignificant

Ninety-eight of the 111 tested hypotheses of equal performance among color treatments could not be rejected at the .05 significance level. For these research hypotheses, which comprised a majority of all the cases, there was insufficient evidence to conclude any difference in the decision maker's performance that could be attributed to the particular color (B/W, CRT, RND, OUT or PAS) of the mixed graphics interface.

Color Factor Significant

These models which preliminary analysis of the data indicated sufficient evidence that treatment differentially affected performance correspond to those

whose entries have an asterisk (*) in Table 4.8. These are listed separately in Table 4.9, Hypotheses Rejected at $\alpha = .05$. There are 13 such cases: females (4), those who have between 3 and less than 5 years college experience (1), education level (2), CRT experience (1) and video game experience (5). Twelve of these had performance measured by quality of decision and one by decision time. These 13 are analyzed in-depth next using Duncan, Tukey-Kramer or the 95% Confidence Interval Graphical approach. Tables 4.10-4.14, constructed by characteristic factor, delineate the results.

Sex. The data did not indicate color affected the performance, whether it was measured as quality of decision or time when the subjects were male. However, the data suggested females were affected by color with respect to quality of decision. Table 4.10, Analysis of Female x Treatment, outlines the results. Since homogeneity of variance was satisfied but the sample size/cell was not equal, Tukey-Kramer was used to detect the differences. Across the top of the columns are listed the color treatments in rank order, the best performance to the worst performance. Down the first column are the performance variables which earlier analyses suggested were affected by the color treatment. The entries in the

Table 4.9

HYPOTHESES REJECTED AT $p \leq .05$

Ho: $\tau_i = 0$

Color Treatment the Significant Factor
p Values

Performance \ Factors	Color & Sex (Female)	Color & Col Ex (3<5 Yr)	Color & Edu Lv	Color & CRT Ex	Color & VGame
QUALITY					
Total	.01				.03
Sum	.02				.05
Scene 1					.01
Scene 2&3	.04				
Post Test	.04		.04		.05
Extrap 1			≥2 & UnG .05		
Extrap >1					.05
Long Mem				.04	
TIME					
Total		.03			

Table 4.10

ANALYSIS of FEMALE vs COLOR TREATMENT

**Performance Not Equal
Across
All Treatments
At $p \leq .05$**

Variances Equal

Color Factor / Performance	CRT Best	B/W	PAS	RND	OUT Worst
QUALITY					
Total (%)	77.6	66.6	66.1	61.2	53.3
Sum (%)	78.6	70.4	68.9	64.0	52.5
Scene 2&3 (%)	78.8	68.8	67.5	64.4	53.3
TUKEY-KRAMER	A	A B	A B	A B	B

Color Factor / Performance	CRT Best	PAS	B/W	RND	OUT Worst
QUALITY					
Post Test(%)	75.7	62.7	58.7	56.4	54.7
TUKEY-KRAMER	A	A	A	A	A

table are the mean scores for each color treatment vs performance variable. The rank order for the treatments and Tukey Kramer groupings were consistent for the three cases measuring Total Quality Score, Sum (Interactive) Quality Score and Scenario 2 & 3 Quality Score. These three performance measures were highly correlated so the results were expected to be consistent. The Tukey-Kramer analysis (α=.05) grouped (CRT B/W PAS RND) as coded by A in Table 4.10 and (B/W PAS RND OUT) as coded by B. The CRT color treatment ranked the best with B/W the second best though analysis did not indicate a significant difference between these two. The CRT treatment yielded a 16.5 per cent increase in Total Score over the B/W treatment, a 11.6 per cent increase in Sum Score and a 14.5 per cent increase in Scene 2&3 Score over the B/W treatment. Overall the CRT treatment yielded a 14.2 per cent increase in decision quality performance when compared to B/W when the decision makers were females.

The consistency was not upheld for the Post Test. B/W and PAS switched ranks, but the others remained the same. The CRT treatment shows a 28.9 per cent increase in Post Test score when compared to the B/W treatment. However, Tukey-Kramer did not detect any significant differences among any of the treatments, which may account

for the ranking inconsistency.

When measuring performance by overall quality of decisions which the Total, Sum and Scene 2&3 quality scores represent, females consistently performed better using a CRT colored mixed graphics display rather then the B/W monochrome intensity pattern one. Though the difference was not statistically significant at $\alpha = .05$, it represented an average of 16.5 per cent in overall (Total) score, 11.6 per cent in interactive (Sum) score, 14.5 per cent in memory interactive (Scene 2&3) score and a 28.9 per cent in long term memory (Post Test) score.

College Experience. Those who had between 3 and less than 5 years college experience showed a preference for specific color treatments when measuring the overall performance time (Table 4.11). Tukey-Kramer (equal variances, unequal cell sizes) suggested no difference in the average total time between RND and B/W coded A in Table 4.11 as well as between B/W, PAS, CRT and OUT coded B in Table 4.11. RND was the best with B/W being second. Though not statistically significant at $\alpha = .05$, this group worked quicker when interfacing with a chromatic display. Unlike the females, the RND schematic ranked first yielding a 2.7 per cent decrease in time over the B/W. The CRT treatment represented a 11.5 increase in time when

Table 4.11

ANALYSIS of COLLEGE EXPERIENCE vs COLOR TREATMENT
(3 < 5 years)

Performance Not Equal
Across
All Treatments
At $p \leq .05$

Variances Equal

Color Factor / Performance	RND Best	B/W	PAS	CRT	OUT Worst
TIME Total (Secs) TUKEY-KRAMER	1826.6 A	1878.0 A B	1934.0 B	2093.5 B	2339.5 B

compared to the B/W.

Education Level. The ANOVA F statistic suggested those who had more than two years of college and were of undergraduate status were affected by the color treatment when they made decisions involving extrapolation from one graph. However, Tukey-Kramer grouped all treatments into one class (coded A in Table 4.12), thus detected no differences among the means. This conflict may be caused by the fact the ANOVA F statistic yielded a p=.05 and Tukey-Kramer was not as sensitive, possibly because of unequal cell sizes to discriminate any differences at .05 significance level. None the less, the CRT cases yielded a 27.9 per cent lower score than the B/W cases while the PAS showed a 5.7 per cent increase over the B/W. Chromatic ranked best.

A 95% confidence interval using the t distribution was constructed around the mean post test scores when considering all levels of education. Since there was insufficient evidence to support the data were skewed rather than normal/bell shaped (Table 4.4) and estimates of the unequal variances were available, the t-distribution was used when building the confidence intervals shown in the second table of Table 4.12. These confidence intervals are plotted in the accompanying

Table 4.12

ANALYSIS of EDUCATION LEVEL vs COLOR TREATMENT

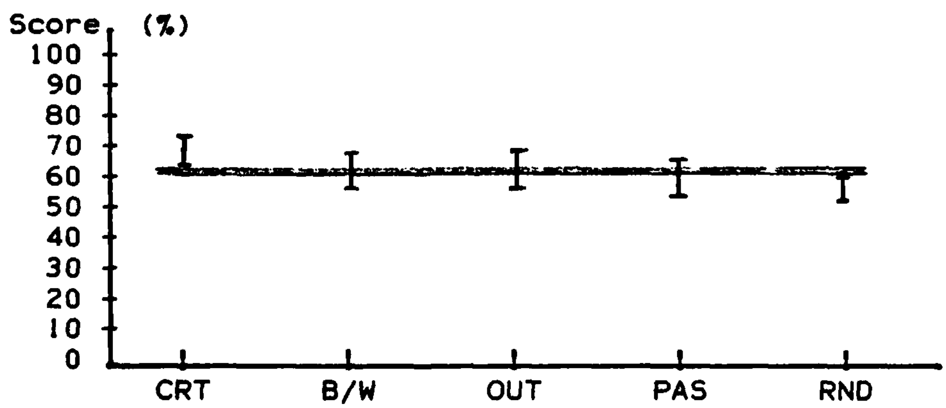
Performance Not Equal
Across
All Treatments
At $p \leq .05$

Variances Equal

Color Factor / Performance	PAS Best	B/W	OUT	CRT	RND Worst
QUALITY Extrap 1 (%) TUKEY-KRAMER	49.6 A	46.9 A	41.1 A	33.8 A	29.0 A

Variances Not Equal

Color factor / Performance	CRT Best	B/W	OUT	PAS	RND Worst
QUALITY Post Test(%) 95% Conf Int TUKEY-KRAMER	67.4 (63,71) A	62.6 (57,67) A B	61.3 (55,67) A B	59.6 (53,65) B	57.7 (52,62) B



graph. The graph indicates that two of the intervals do not intersect: CRT (63.5, 71.5) and RND (52.8 and 62.6). These two may belong in different groups. If Tukey-Kramer was used, its grouping confirms this hypothesis. It grouped CRT B/W and OUT (coded A in Table 4.12) as not significantly (α =.05) different from each other and neither were B/W OUT PAS and RND, coded B in Table 4.12. From the ranking of means, CRT color treatment was the best, with B/W second when measuring performance by a long term memory (Post Test) score. Examination of the confidence intervals indicate though the CRT and B/W treatments were not significantly different, the CRT represented approximately a 4 per cent increase in performance over B/W.

CRT Experience. Duncan's Multiple Range test was applied to the data categorized by the Color x CRT Experience vs Long Term Memory Quality Score (Table 4.13). It detected differences among the means and divided the 5 equal size groups into 2 categories. The first included (CRT B/W RND) and the second (B/W RND OUT PAS). CRT ranked first, with B/W in second place, with CRT outscoring B/W by 9.8 per cent. These two, like all previous cases were not statistically significant from each other.

Table 4.13

ANALYSIS of CRT EXPERIENCE vs COLOR TREATMENT

Performance Not Equal
Across
All Treatments
At $p \leq .05$

Variances Equal

Color Factor / Performance	CRT Best	B/W	RND	OUT	PAS Worst
QUALITY Long Mem (%) Duncan Range	69.4 A	63.2 A B	61.1 A B	59.0 B	58.3 B

Video Game Experience. After accounting for the effects due to video game experience, the color of the graphics display became significant when performance was measured by Total Score, Sum Score, Post Test Score, Extrapolation Using >1 Graph and Scenario 1 Score. The first four were further analyzed using Duncan's Multiple Range since there was insufficient evidence to reject homoskedasticity or additivity. The 95% confidence interval graphical method was applied for the Color x Video Game vs Scenario 1 score based on Levene's test supporting unequal variances. The results are delineated in Table 4.14, Analysis of Video Game x Treatment.

Controlling $\alpha=.05$, Duncan's analysis indicated the mean total score was the same for the CRT, B/W, PAS and OUT color treatments as well as for PAS, OUT and RND. CRT ranked best with B/W in second place with no significant difference between them. The CRT treatment only yielded a .7 per cent increase in total performance over B/W.

The ranking of the sum score is not consistent with the total despite the high correlation between the two performance measures. The order is PAS, B/W, CRT, OUT and RND with the first four not statistically different from each other. The last three were not either in accordance with Duncan. PAS yielded the best average score while B/W

Table 4.14

ANALYSIS of VIDEO GAME EXPERIENCE vs COLOR TREATMENT

Performance Not Equal
Across
All Treatments
At p≤.05

Variances Equal

Color Factor / Performance	CRT Best	B/W	RND	OUT	PAS Worst
QUALITY Total (%) Duncan Range	68.9 A	68.4 A	67.8 A B	65.3 A B	61.6 B

Color Factor / Performance	PAS Best	B/W	CRT	OUT	RND Worst
QUALITY Sum (%) Duncan Range	72.6 A	71.6 A	69.8 A B	67.8 A B	63.8 B

Color Factor / Performance	CRT Best	B/W	OUT	PAS	RND Worst
QUALITY Post Test(%) Duncan Range	67.5 A	62.6 A B	61.3 A B	59.6 B	57.7 B

Color Factor / Performance	B/W Best	PAS	OUT	CRT	RND Worst
QUALITY Extrap >1(%) Duncan Range	72.5 A	70.3 A	61.4 A	56.9 A	53.3 A

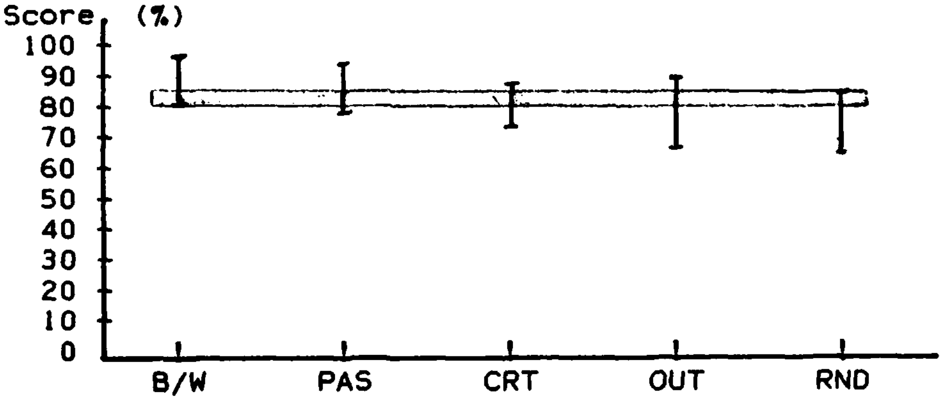
Table 4.14 (Cont.)

ANALYSIS of VIDEO GAME EXPERIENCE vs COLOR TREATMENT

Performance Not Equal
Across
All Treatments
At $p \leq .05$

Variances Not Equal

Color Factor / Performance	B/W Best	PAS	CRT	OUT	RND Worst
QUALITY					
Scene 1 (%)	87.9	86.3	81.7	79.5	75.7
95% Conf Int	(80,95)	(78,94)	(74,88)	(69,90)	(66,85)
TUKEY-KRAMER	A	A	A	A	A



rated second. PAS represented only a 1.3 per cent increase in interactive (Sum) score over B/W, with CRT showing only a 2.5 per cent decrease in performance with respect to B/W.

The average post test scores were also separated into two intersecting groups. Duncan included CRT, B/W and OUT in the better group and B/W, OUT, PAS and RND in the second one. As in the Total Score performance, CRT ranked highest and was not statistically different from B/W which came in second best for the Post Test Score. CRT yielded a 7.8 per cent increase in the long term memory (Post Test) score over B/W.

After accounting for the effects of video game experience, color was considered significant based on the ANOVA F at p=.05 when performance was measured by the Extrapolation Using >1 Graph. However Duncan did not discriminate any significant differences among any of the treatment means. However B/W ranked first, representing an increase in the Extrapolation >1 Graph score of 3.0 per cent over PAS, 15.3 per cent over OUT and 21.5 per cent over CRT.

When analyzing the graphed 95% confidence intervals computed for the Scenario 1 Score means, there was not an obvious split among groups. The (80.9, 85.2) score

interval was contained in all of them. Also, if the Duncan and Tukey-Kramer tests were applied as estimates of the differences, neither reported significant differences among any of the Scenario 1 Score means, which supports the graphical analysis. Similar the the Extrapolation >1 Graph score just discussed, achromatic outscored chromatic with the ranking order of B/W, PAS, CRT, OUT and RND. B/W outscored PAS by about 1 per cent and CRT by approximately by 4 per cent.

After accounting for the variability in performance caused by video game experience, when in-depth analysis indicated difference among treatments, chromatic displays consistently ranked highest with B/W in second place. The difference was always statistically insignificant at $\alpha = .05$, with the average increase in score only being 4.9 per cent. CRT showed an increase of .7 per cent in overall (Total) decision quality over B/W; PAS a 1.3 per cent increase in interactive (Sum) decision quality over B/W, and CRT a 7.8 per cent increase in long term memory (Post Test) score over B/W. Though the differences were never significant and the increase in decision quality small, as measured by scores comprised of all types of decisions, performance was best when chromatic displays were the interfaces. The CRT ranked first twice and PAS ranking

first once. The B/W monochrome intensity pattern consistently was second in performance.

Conclusions

The research question addressed in this chapter was how the attribute of color, when used functionally in computer generated mixed graphics displays affects the performance of a decision maker who has an academic engineering background. Five color treatments were used in a laboratory experiment. One treatment was achromatic consisting of black and white computer generated intensity patterns. The other four were chromatic, representing possible variations of CRT color outputs. These treatments were functionally incorporated into hardcopy mixed graphics displays used by 120 subjects enrolled in the College of Engineering, ASU to complete a one hour managerial planning deterministic decision making case study. Potential variations in data due to seven different characteristics of the subjects were controlled prior to analysis of the research question. Performance was measured as a function of time (3 levels) and as a function of decision quality (12 levels). As a result, 111 hypotheses and ANOVA models were formulated and validated for rigorous analysis of the effects of color on performance. These 111 cases are shown in Table 4.8,

Analysis of Main Effect of Color.

Of the 111 tests, 98 resulted in insufficient evidence to conclude that a significant difference in performance occurred among any of the five treatments. Even after the affects of characteristic factors such as age, sex, college level and experience, attitude toward the computer, computer experience and video game experience were taken into account in the variation of the mean performance, the color attribute did not test statistically significant at the .05 significance level.

For those 13 cases listed in Table 4.9, Hypotheses Rejected at $\alpha = .05$, preliminary analysis indicated the data contained evidence that supported that the attribute of color in the mixed graphics display did affect the performance of decision makers. For 11 of these, shown in Table 4.15, Ranking of Treatments, in-depth analyses further supported differences in performance attributable to color treatment. The table lists the rank order across the top, from best performance to worst. The first column lists the characteristic factors and performance measures of the hypothesis of equal mean performance which were rejected at the .05 significance level. The entries in the table are the codes representing the color treatment of the mixed graphics display. They are ranked according

Table 4.15

RANKING of TREATMENTS
by
PERFORMANCE

Groupings of Duncan/Tukey-Kramer
$\alpha = .05$

Color Ranking / Characteristic vs Perform	Best 1	2	3	4	Worst 5
Female vs Total Score	CRT	B/W	PAS	RND	OUT
Female vs Sum Score	CRT	B/W	PAS	RND	OUT
Female vs Scene 2&3	CRT	B/W	PAS	RND	OUT
Col Ex vs Total Time	RND	B/W	PAS	CRT	OUT
Edu Lv vs Post Test	CRT	B/W	OUT	PAS	RND
CRT Ex vs Long Mem	CRT	B/W	RND	OUT	PAS
V Game vs Total Score	CRT	B/W	PAS	OUT	RND
V Game vs Sum Score	PAS	B/W	CRT	OUT	RND
V Game vs Post Test	CRT	B/W	OUT	PAS	RND

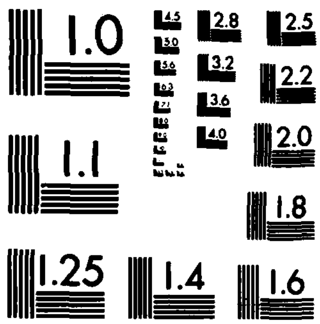
to the top line. The underlining of each group represents the grouping of equal means suggested by Duncan or Tukey-Kramer, as discussed earlier.

Analyzing the content of the table suggests several trends. If there was evidence of significance, the means were all divided into only two intersecting subgroups. The top ranking treatment was consistently chromatic, with the second always the monochrome intensity pattern (B/W). Except for the Color x Video Game vs Total Score or Sum Score, the monochrome pattern was common to all subgroups, consistently being the second best in the better group and top in the second group. Of the chromatic treatments, the CRT colors (CRT) ranked the highest most often while the pastel colors (PAS) were middle of the road and the random CRT (RND) and outline (OUT) ranked at the worst end. For the seven cases where CRT outscored the B/W, all which were a function of quality of decision, the average increase in performance was 9.2 per cent. However in the five cases where B/W outscored CRT, the average decrease in score was about the same, 13.9 per cent. Of particular note is that model which had decision time as the performance measure. B/W decreased decision time over CRT by 11.5 per cent.

Overall, based on the results of 102 of 111 tested

AD-A145 561 EFFECTS OF ALTERNATIVE CHROMATIC MIXED GRAPHICS DISPLAYS IN DECISION SUPPORT SYSTEMS(U) AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH M S MCCULLY MAY 84 3/4

UNCLASSIFIED AFIT/CI/NR-84-46D F/G 9/2 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

hypotheses, the attribute of color when functionally integrated in a mixed graphics display did not significantly affect performance in either direction. Quality of decision was not improved or degraded at the .05 significance level, nor was decision time. The data did not contain sufficient evidence to conclude the chromatic colors have a unique quality for functionally coding information in a DSS over a monochrome intensity pattern. However, if color is incorporated functionally into a DSS whose interface is hardcopy, the CRT color set appears to be the best in general. However, it should not be assumed the performance of the decision maker with an academic engineering background will be significantly affected.

Conclusions and generalizations based on these results and the literature as well as recommendations for further research in this area are offered in the next chapter.

V. CONCLUSIONS AND RECOMMENDATIONS

Introduction

Computer generated graphics supporting the option of color have become the heart of many DSS in the industrial, business and military sectors. However knowledge of how to integrate this color technology to realize maximum benefit to the performance of the decision maker has not kept pace with the technological advances of the software and hardware supporting it. The design objective of a DSS from a manager's perspective is to discover the optimum match between the DSS and the decision maker that allows for more efficient and effective performance. The primary concern of the research reported here was to investigate the relative effectiveness on performance when integrating color functionally into a DSS whose decision maker has an academic engineering background, solves structured problems based on deterministic data presented in a mixed graphics form on computer generated hardcopy in a managerial role of control of resources.

The MIS literature advocates empirical research is desperately needed in this area and recommends the small study (Lucas, 1981) laboratory (Dickson et al., 1977) approach. Since it was determined the objective of this study could be satisfied within the framework of a

laboratory setting, an experiment was designed using this recommended approach. The key factors of a DSS definition as given by Mason and Mitroff (1973) determined the elements of the experiment. The performance measured were two resources most coveted by decision makers: time and quality of decision. These two variables were subdivided into 15 categories for more rigorous analysis of the research question. Three categories were a function of time; 12 were a function of quality of decision. Seven characteristics of decision makers that may have influenced performance were included in the design. These included age, sex, college level and experience, attitude toward the computer, and experience with computers and video games. The case study used as the instrument represented a general class of structured problem of managerial resource control. The mixed graphics form was bar charts that are widely used on existing graphical DSS. Five alternatives of colors were functionally integrated into the graphics for comparison. Four of these alternatives were chromatic and included: the colors preset on a majority of CRTs (CRT), these same colors but in outline as opposed to filled form (OUT), a set of colors randomly selected from all colors available on a majority of CRTs (RND) and a set of randomly selected

pastel colors (PAS). The fifth alternative was achromatic or monochrome and was comprised of randomly selected intensity patterns (B/W) available in many commercial computer graphics software packages. The man-machine interface employed was computer generated hardcopy.

Based on the overall objective of this research and the variables designed into the experiment as discussed above, hypotheses were formalized. The data collected from the performance of 120 subjects were used to define the appropriate mathematical models whose derived statistics were used to rigorously test the stated hypotheses.

This chapter initially reiterates the results of examining the problem statement via analyses of the 111 hypotheses. This is accomplished by offering conclusions and summaries first by individual performance variable, then by performance variable type (decision quality and time) and finally by generalizing to the overall problem statement and objectives of this research. Recommendations based on recognized limitations of this research and direction from as well as voids in the MIS literature finish this chapter and conclude this research.

Conclusions and Summary

The following conclusions and summaries are based on

the results of rigorous analyses of the appropriate statistics derived from the 2-factor fixed effect ANOVA model. The particular statistics used and their justification based on the characteristics of the model as well as the data are detailed in Chapter IV, Analysis and Results. The final results are consolidated in Table 5.1, Final Results of Analysis of Color. The entries in this table are the p values of the models whose performance variable starts the row and whose color and characteristic factors head the column. The p value entries are included only for those models which in-depth analyses via Duncan, Tukey-Kramer or graphical inspection indicated sufficient evidence to reject the null hypothesis that performance was not affected by the color alternative of the mixed graphics display at the .05 significance level.

The conclusions and summaries are presented as follows. First those models whose performance was a function of decision quality are discussed with a general summary offered. Next the models whose performance was a function of time are presented then summarized. Overall conclusions and generalizations close this section.

<u>Decision Quality Score</u>

<u>Total Score</u>. Total score represented the overall decision quality. There was insufficient evidence to

Table 5.1

FINAL RESULTS of ANALYSIS of COLOR

p Values of models
Equal Performance Rejected at $\alpha = .05$

Performance \ Factors	Color & Age	Color & Sex	Color & Col Ex	Color & Edu Lv	Color & CRT Ex	Color & Attit	Color & VGame
QUALITY							
Total		.01*					.03
Sum		.02*					.05
Scene 1							
Scene 2&3		.04*					
Post Test				.04			.05
D Read >1							
Extrap 1							
Extrap >1							
Short Mem							
Long Mem					.04		
Long Mem Extrap							
TIME							
Total			.03#				
No Memory							
Memory							

* Females Only
# $3 \leq 5$ Yrs Col Exp Only

conclude that color influenced the overall decision quality for all models except two. Females' overall decision quality score ranked best to worst when interfacing with CRT, B/W, PAS, RND and OUT respectively. In-depth analysis showed the first four not significantly different as well as the last four. When video game experience was washed out, the ranking was similar to females: CRT, B/W, PAS, OUT and RND. The first four were insignificant as were the last three.

Overall total decision quality was not affected by color of the graphics display. When there was evidence to suggest a significant difference, the rankings were similar: CRT, B/W and PAS. However, there was insufficient evidence to suggest chromatic (CRT) which ranked best was significantly better than monochrome (B/W), ranked second. The actual average increase of CRT score over B/W was 8.6 per cent.

Sum. Sum score represented quality of decisions based on information which was either currently or previously available to the decision maker. As with the total score, for all but the same two models, there was insufficient evidence to conclude color significantly affected performance. This was expected since sum score is a subset of total score. Females' performance were

consistent in its differential ranking: CRT, B/W, PAS, RND and OUT. In-depth analysis grouped the first four as a set and the last four as a set. Chromatic (CRT) was not significantly different from monochrome (B/W) though it outscored monochrome by 11.6 per cent.

After washing out the variability in performance due to video game experience, decision makers sum score ranked best to worst for PAS, B/W, CRT, OUT and RND. In-depth analyses grouped the first four as insignificant from one another and the last three as insignificant from each other. Chromatic (PAS) was not significantly different from monochrome (B/W) though it outscored it by 1.3 per cent. Unlike all above discussions, the PAS alternative ranked first. CRT, which previously ranked first was outscored by B/W by 2.5 per cent, which was not statistically significant.

Scene 1. Scene 1 quality score represented decisions in which no memory was required. In-depth analysis indicated insufficient evidence to concluded significant differential affect on no memory decision quality by any of the treatments.

Scene 2 & 3. Scene 2 & 3 quality score represented decisions requiring information that was present for review as well as memory of what was presented or decided

during the previous scenario(s). All models except female performance showed indifference to the color of the graphics displays. Females performed these types of decisions best to worst when interfacing with CRT, B/W, PAS, RND and OUT respectively. This ranking, which in-depth analysis showed the first four insignificant from each other as well as the last four, was consistent with previous rankings by female performance. Chromatic (CRT) outscored monochrome by 14.5 per cent which was not statistically significant.

Post Test. Post test quality of decision represented decisions for which none of the graphics displays, problem facts or previous information were available and the session was not timed. It required memory and/or manipulation of previously given information or decisions. After in-depth analyses were accomplished, only two of the eight models contained evidence to suggest the color of the graphics display affected post test decision performance: those which accounted for the variability due to education level and for video game experience. Both models were consistent with each other in their respective rankings of color based on post test performance: CRT, B/W, OUT, PAS and RND. The first three were grouped as insignificant as were the last four.

After the variability in performance attributed to education level or video game experience was accounted for, chromatic (CRT) outscored monochrome (B/W) by approximately 4 per cent and 7.8 per cent respectively. This represented an average increase of approximately 5.9 per cent in post test score using CRT over B/W. However this difference was not statistically significant at the .05 level. The other six models whose performance measured post test quality score contained insufficient evidence to support any of the five alternatives were more effective than any of the others.

D Read >1. Direct read from more than one graph represented decisions requiring the processing of information from at least two graphs that were available for review by the decision maker. None of the eight models investigated were supported by data which suggested color differentially affected this type of performance.

Extrap 1. Extrapolation from one graph represented that type of decision which required manipulation of the information from one graphics display which was available for review by the decision maker. In-depth analysis indicated insufficient evidence to conclude color differentially affected this type of decision performance.

Extrap >1. Extrapolation from more than one graph

was comprised of decisions which required computation of information extracted from two or more graphs but did not require memory. In-depth analyses resulted in failing to reject that performance based on all five treatments was the same.

Short Mem. Short term memory score represented decision quality which mandated memory of previously presented information and/or decisions. Computation was not required. The case study, current information and graphics displays were available for review by the decision maker. For all models under consideration, there was insufficient evidence to conclude that a particular color differentially affected a decision maker's short term memory quality of decision performance.

Long Mem. Long term memory decisions required memory of previously given information or stated decisions. No computations were necessary. Unlike the short term memory decisions, neither the problem facts or any graphics displays were available for review. This is the first of the decision types that was not timed. Except for the model which accounted for the variation in performance due to CRT experience, there was insufficient evidence to conclude the color of the graphics display significantly influenced performance which required long term memory.

As these decisions were part of the last set made by the subject, the same results can be generalized to the affects the color alternatives had on maintaining the attention of the decision maker.

In-depth analysis showed, that after accounting for the effects on performance attributed to CRT experience, the color of the graphics display did affect long term memory performance. Performance was best to worst when the displays were CRT, B/W, RND, OUT and PAS respectively. The first three were grouped as being insignificant from each other as were the last four. Chromatic (CRT) outscored monochrome (B/W) by 9.8 per cent, which statistically did not represent a significant difference.

Long Mem Extrap. Long term memory extrapolation type decisions represented that class of decisions which requires manipulation of information or decisions previously given. Neither the basic facts or any of the graphics displays were accessible to the decision maker. These types of decisions were not timed and represented the end of the decision making session. Across the board, there was insufficient evidence to reject the hypothesis that any of the five color alternatives varied from another when quality of decision was measured by long term memory performance. This may also indicate that none of

the alternatives were more influential on maintaining subject attention better than another set.

Summary Based on Decision Quality

Of the 86 models that were examined whose performance was a function of quality of decision, the majority (78) of them resulted in insufficient evidence to conclude the color attribute of the mixed graphics display significantly affected performance at the .05 significance level. These results are inconsistent with the theoretical and subject literature, which purport color when integrated into information display allows for better decisions and memory. *However, these findings are* concurrent with the objective literature which was summarized by stating chromatic color does not appear to have a unique quality for coding information over other achromatic codes (Christ and Corso, 1983). For some tasks, it may be beneficial; in others it may be detrimental. In this research, chromatic color did not influence performance in a way different than the achromatic code for a majority of the models. Where it did, the influence was always in the positive direction as it consistently ranked better than achromatic, though never with an interval which was statistically significant at the .05 significance level.

Unlike any other study reported in the literature, this research employed several chromatic alternatives in the experimental design. This was done to investigate if it was just the presence of color that affected performance or a particular set. As discussed previously, decision quality performance was not affected differently by any of the five alternatives. Based on these results, it can be concluded that despite the color set used, performance was not affected differently than when a monochrome intensity pattern was used.

After examining the eight models whose data contained sufficient evidence which suggested chromatic color affected decision quality performance, a pattern was detected. In seven of the eight models, the CRT alternative correlated with the best performance, giving an average of approximately 9.2 per cent increase in decision quality score over the second place monochrome intensity pattern. This difference was not statistically significant. For the eighth case, PAS ranked best, showing a 1.3 per cent increase over monochrome. Based on these results it is concluded that if a color alternative is chosen, the data suggest the colors preset on the majority of CRTs be the ones used to maximize performance. However, since not all nor a random sample from all color

alternatives available for graphics displays were investigated, the generalization of this conclusion is limited. When choosing among the four chromatic alternatives, quality of performance was most often positively affected by the colors currently preset on a majority of CRTs supporting color: green, red, yellow, blue, dark blue, cyan, black and white.

Decision Time

Total. Total or interactive time represented the time (in seconds) needed to make those decisions where the basic problem facts and graphics displays were available for review. Eight of the nine models' data considered contained insufficient evidence to conclude chromatic displays differentially affected the over all time required to make decisions. For those decisions makers who had at least three but less than five years college experience, decision time was shortest when the visual display was functionally colored with the RND alternative, with B/W, PAS, CRT and OUT being the remaining ranking order. In-depth analysis showed RND outscored B/W by 2.7 per cent, while B/W shortened total time by 11.5 per cent over CRT. RND and B/W were not significantly different from each other, nor were B/W, PAS, CRT and OUT.

No Memory. No memory time was the time (in seconds)

elapsed when formulating decisions which required no memory but extraction of information currently available to the decision maker. Results from analyses of the no memory time data were consistent across every case. Functional chromatic information code did not affect decision time when compared to monochrome intensity pattern.

Memory. Memory time was the time (in seconds) spent by the decision maker when making decisions which required memory for some, intermingled with others which did not. The problem facts and current graphics displays were available for consulting, but the necessary information was not always currently available. Analysis overwhelmingly concluded the color of the graphics display did not significantly affect memory time.

Summary Based on Decision Time

Of the 25 models that were examined whose performance was a function of decision time, all but one resulted in insufficient evidence to conclude the color attribute of the mixed graphics display significantly affected decision time at the .05 significance level. Like the results based on decision quality, these are inconsistent with the theoretical and subjective literature which contend color significantly speeds the information processing thus

decision time. The single case in which in-depth analyses indicated a difference on performance attributed to color, the best chromatic set decreased decision time by only 2.7 per cent, but monochrome showed a decrease in decision time by 11.5 per cent over the suggested chromatic scheme of CRT based on decision time performance. Based on this, it is concluded that chromatic code does not positively influence a decision maker's performance time, whether this time is measured as an overall time or time required to make no memory or a mix of no memory/memory decisions.

Overall Conclusions Generalized to Problem Statement

The primary objective of this research was to investigate the relative effectiveness on performance when functionally integrating color into a DSS whose decision maker has an academic engineering background, solves structured problems based on deterministic data presented in a mixed graphics form on a computer generated hardcopy in a managerial role of control of resources.

Overall, based on the results of 102 of 111 tested hypotheses, the attribute of color when functionally integrated in a mixed graphics hardcopy display did not significantly affect performance in either direction of a decision maker with an engineering background. Quality of decision was not improved or degraded significantly at the

.05 significance level, nor was decision time.

In the nine models where the data contained evidence that the color alternative affected performance, the CRT alternative ranked best in seven instances, all which were a function of decision quality. The average improvement was approximately 9.2 per cent. Specifically, females performed best when interfacing with the CRT alternative, but it was never a significant improvement over the monochrome (B/W) intensity pattern.

In the other two models where a difference in performance was suggested, a chromatic code ranked first, but it varied (RND, PAS) and was not CRT. The improvement was never significant with respect to the monochrome. Monochrome intensity pattern always ranked second. It showed a 11.5 per cent improvement in performance time and only a 2.5 per cent improvement for sum quality score over the CRT alternative.

In general, these results are concurrent with other empirical studies on the effects of color when integrated into visual displays. The data do not contain sufficient evidence to conclude the chromatic colors have a unique quality for functionally coding information in DSS over a monochrome intensity pattern. However if color is functionally added to a hardcopy mixed graphics display,

the preset colors available on a majority of the current CRTs appear to be the best overall. However, the manager should not assume increased efficiency in a decision maker's performance because of the functional color attribute.

Areas for Further Research

The DSS is composed of five key factors: a person of a certain psychological type, a class of problems, data of a certain form, a particular man-machine interface and an organizational setting. This research is limited in scope to the decision maker with an academic engineering background, who solves deterministic problems whose data is presented in mixed graphics form on a hardcopy interface when assuming a managerial control of resources role. Changing any one factor opens the door for further research. Based on the recommendations of the MIS literature that the approach to research be that of smaller studies with emphasis on the laboratory setting, the users of DSS supporting color, the technology of the color itself and the limitations of this research, the following recommendations are made for further research.

Recommendations

1. The CRT is gaining as the primary man-machine interface in the DSS with the user friendly, available and

affordable computer system technology. Color most likely will become a standard option at a cost not much different from that of monochrome. How does color affect the performance of the decision maker whose interface is primarily the CRT? If differences exist, are they the same for either type of monochrome (black/white and green phosphorous) CRT interface? Performance as a function of memory will be an issue since once the screen is erased for the next graphics output, the first is not available as in hardcopy. Performance as a function of time may also be a key issue. If memory is enhanced by color, fewer graphics will have to be recalled to the screen for reexamination, thus the decision time shortened. With the dark background of most CRTs as opposed to the white of the hardcopy, and the horizontal eye level of the CRT versus the 45 degree to 80 degree angle for hardcopy on the desk, the impact of color may vary with interface.

2. Though quality of decision and time are the two most coveted resources of most decision makers, others have high utility. Does color in the DSS make the job more exciting thus give the decision maker added motivation? The subjective literature purports color commands a longer attention span. Does it? The subjective literature also claims color causes less

fatigue. Does color reduce fatigue in a short interval or over an extended period of time? Does color affect the performance differently than monochrome irrespective of the fatigue of the decision maker prior to interfacing with the DSS?

3. If the amount of viewing time for the mixed graphics display is limited, in what, if any, situations does color aid the decision maker? This study was decision maker paced, but not all situations are.

4. Christ (1977) criticized that all the research on color to date was accomplished in an isolated laboratory setting in which the inexperienced subject dedicated total attention to performing simple discrete task(s). This research employed a complex task, similar to one which most subjects had previously performed on their own, in their workplace or in an academic setting. However, unlike the decision maker's environment, the task was accomplished in a testing, uninterrupted classroom. The environment allowed total energies to be dedicated to the task. Recommend a similar study be accomplished which includes in its design interruptions that simulate the decision maker's workforce environment.

5. This study sampled a population within an academic environment. The thought processes and

strategies when solving the case study in a classroom testing environment may be different from those in the engineering workforce. Recommend analysis of the affects of color on performance of decision makers currently in the workplace who are not involved in academic training. Special attention in the analyses should be given to the different ages, experience levels and strata of management as each may use the same DSS from a different perspective.

6. Are decision makers with academic engineering backgrounds affected differently by color than decision makers with nonengineering backgrounds? Those decision makers with an artistic, architecture, educational etc background may be stimulated differently by color displays. Research to date is void in this area.

7. This research did not follow up on subjects to test memory over an extended period. Often managers make decisions based on information which they reviewed days or weeks earlier. Does color affect memory over an extended period of time?

8. Not all problems faced by decision makers using a DSS fall into the general class of problems represented by the inventory control problem whose data were or were considered deterministic. Does color relay information differently for different types of problems? The

experimental literature reports most decision makers remove the stochastic nature from their data prior to their use. Does information processed in color affect this process?

9. Research on the effects of ambient lighting when using the CRT, with recommendations on its levels has been reported in the literature. However, these studies have been mainly directed to CRT visual displays supporting text, with emphasis on monochrome displays. Do the recommendations for the ambient lighting remain the same for CRT displays supporting color graphics displays?

10. Vertical bar mixed graphics displays as used in this research are one of the most widely forms of displays used. However, they are not the only type. Does color, when functionally added to other types of graphics displays such as line graphs, pie charts, horizontal bar charts, etc affect the performance differently than the monochrome patterns?

11. Often information on which decisions are based is presented in a meeting environment with the interface being viewgraphs or slides, usually read at a distance. As a result of the technology, the viewgraphs and/or slides are images of the CRT displays. How does color, when incorporated in this type of decision maker's

interface, affect the performance of a decision maker in the short and long term?

12. Theoretically, color should enhance performance. The problem may lie in the fact it has not yet been discovered how or what colors to use that will take maximum advantage of how the brain processes the information code. What colors? How many colors? How much of each color? are questions which have not been rigorously addressed and remain unanswered.

Conclusions

There continues to be a large gap of quantitative information regarding how to integrate color in a DSS so it benefits the user. This reported study attempted to narrow this gap of knowledge. Research efforts like this and those suggested above may lead to discovering a way of allowing color to enhance the effectiveness of a DSS.

BIBLIOGRAPHY

Air University compendium of research topics, 1982-83, Vol. 1. Air University (ATC), Maxwell Air Force Base, Alabama, July 1982.

Barden, W. High-resolution graphics system. Popular Computing, 1982, 1, 116-120.

Bariff, M. & Lusk, E. Cognitive and personality tests for the design of management information systems. Systems Management, April 1977, 23, 820-829.

Barker, E. & Krebs, M. Color effects on human performance: An annotated bibliography. Technical Report ONR-CR213-136-1f. Minneapolis, Minnesota: Honeywell, Inc., April 1977.

Barr, A. (editor). SAS user's guide, 1979 edition. North Carolina: SAS Institute Inc., 1979.

Belie, R. & Rapagnani, N. Color Computer Graphics - a design and analysis potential scarcely tapped. Astronautics and Aeronautics, June 1981, 19, 41-48.

Benbasat, I. & Taylor, R. The impact of cognitive styles on information system design. Management Information Systems Quarterly, June 1978, 2, 43-53.

Benbasat, J. & Schroeder, R. An experimental investigation of some MIS design variables. Management Information Systems Quarterly, March 1977, 1, 37-49.

Benson, T. Graphics capabilities for the small business. Interface Age, February 1982, 7, 66-69+.

Berenson, M., Levine, D. & Goldstein, M. Intermediate statistical methods and applications: A computer package approach. New Jersey: Prentice Hall, Inc., 1983.

Bergman, H. & Duijnhouwer, F. Recognition of VDU presented colors by color defective observers. Proceeding of the Human Factors Society. Los Angeles, 1980, 611-615.

Berk, T., Brownston, L. & Kaufman, A. A new color-naming system for graphics languages. IEEE Computer Graphics and Applications, May 1982, 2, 37-42.

Biggs, L. G. Seven steps for effective graphics presentation. Computerworld, August 31, 1981, 15, Sr/5+.

Bork, A. Computer graphics in learning. Journal of College Science Teaching, January 1980, 1, 141-149.

Box, G., Hunter, W. & Hunter J. Statistics for Experimentation. New York: John Wiley and Sons, Inc., 1978.

Brooks, R. Search time and color coding. Psychonomic Science, 1965, 2, 281-282.

Brown, M. & Forsythe, A. The small sample behavior of some statistics which test the equality of several means. Technometrics, February 1974, 16, 129-132.

Cahill, M. & Carter, R., Jr. Color code size for searching displays of different density. Human Factors, 1976, 18, 273-280.

Calev, F. Enhancing comprehension with color. Mini-Micro Systems, August, 1981, 14, 139-140.

Carlson, W., Parent, R. & Csuri, C. The use of color and 3-d temporal and spatial data management techniques in computer-aided design. Proceedings, 16th Design Automation Conference, June 1979, 32-38.

Catalano, F. Color graphics vendors set new price lows. Mini-Micro Systems, February, 1982, 15, 109-114.

Chard, A. Computer display - a clear view of present methods. Process Engineering, September 1979, 116-117.

Chase, W. Evaluation of several TV display systems configurations for visual simulation of the landing approach. IEEE Transactions on Man-Machine Systems, 1970, MMS-11, 140-149.

Chervany, N. & Dickson, G. An experimental evaluation of information overload in a production environment. Management Science, October, 1974, 20, 1335-1344.

Christ, R. Review and analysis of color coding research for visual displays. Human Factors, 1975, 17, 542-570.

Christ, R. Four years of color research for visual displays. Proceedings of The Human Factors Society, 21st annual meeting. San Francisco, California, October, 1977, 319-321.

Christ, R. & Corso, G. The effects of extended practice on the evaluation of visual display codes. Human Factors, February 1983, 25, 71-84.

Christner, C. & Ray, H. An evaluation of selected combinations of target and background coding on map-reading performance - experiment V. Human Factors, 1961, 3, 131-146.

Chute, A. Effect of color and monochrome versions of a film on incidental and task relevant learning. Educational Communication and Technology, Spring 1980, 28, 10-18.

The color impact in graphic displays. Optical Spectra, April 1980, 14 46-53.

Cook, T. Colour coding: A review of the literature. NTIS No. AD A001555, November 1974.

Conover, W. Practical nonparametric statistics. New York: John Wiley and Sons, Inc., 1971.

Cosky, M. Computer-based instruction and cognitive styles: Do they make a difference? October 1980, (ED 201 299).

Coulter, N. Software and cognitive psychology. IEEE Transactions on Software Engineering, March 1983, SE-9, 166-171.

Crawford, A. Interactive computer graphics for simulation displays. Proceeding of the Human Factors Society, 21st annual meeting. San Francisco, California, October 1977, 14-17.

CRT graphics terminal market to exceed $1.4 billion by 1986. Interface Age, February 1982, 7, 19.

DeMars, S. Human factors considerations for the use of color in display systems. NASA-TM-X-72196, TR-1329, January 1975. NTIS N75-14446.

Denbrook, C. Artistry in layers: Generating color transparencies. Computer Design, July 1982, 21, 155-163.

DI-3000 user manual. Colorado: Precision Visuals, Inc., 1979.

Dickson, G., Senn, J. & Chervany, N. Research in management information systems. Management Science, May 1977, 23, 913-923.

Dixon, W. (editor). BMDP statistical software 1981. Los Angeles: University of California Press, 1981.

Dooley, R. & Harkins, L. Functional and attention-getting effects of color on graphics communications. Perceptual and Motor Skills, 1970, 31, 851-854.

Durrett, H. Color display systems: The state of the art. Behavior Research Methods and Instrumentation, 1979, 11, 127-130.

Durrett, J. & Trezona, J. How to use color displays effectively, April 1982, BYTE 7, 50-53.

Durrett, J. & Zwiener, C. A microcomputer-based laboratory: Real world experience. Behavior Research Methods and Instrumentation, April 1981, 13, 209-212.

Durrett, J., Zwiener, C. & Freund, R. The human factors of color display-based instruction. Proceedings of the National Educational Computing Conference, 1981, 284-286, (ERIC 207 526).

Dvorine, I. Color perception testing charts (second edition). Baltimore, Maryland: Scientific Publishing Company, 1944.

Eddy, D. Application of color displays to problems in the undergraduate laboratory. Behavior Research Methods and Instrumentation, April 1979, 11, 131-134.

Ford, W. Tutorials for the first time computer user. IEEE Transactions Professional Communication, March 1981, vol. PC-24, #1, 30-37.

Francis, E. Using charts to improve profits. New Jersey: Prentice-Hall, Inc., 1962.

Friend, D. Color graphics information systems boost productivity. Mini-Micro Systems, May 1980, 13, 181-182+.

Furst, A. Raster refresh gaining a role in color graphics. Electronic Business, July 1982, 8, 92-94.

Gerrity, T. Design of man-machine decision systems: An application to portfolio management. Sloan Management Review, Winter 1971, 12, 59-75.

Grandjean, E., Hunting, W. & Pidermann, M. VDT workstation design: Preferred sittings and their effects. Human Factors, April 1983, 25, 161-175.

Green, B. & Anderson, L. Color coding in a visual search task. Journal of Experimental Psychology, 1956, 51, 19-24.

Haber, R. & Wilkinson, L. Perceptual components of computer displays. IEEE Computer Graphics and Applications, May 1982, 2, 23-26.

Hallworth, H. & Brebner, A. Computer assisted instruction in schools: Achievement, present developments, and projections for the future. Alberta, Canada: University of Calgary, June 1980, (ERIC 200 187).

Hanson, D. Introduction to the special issue on the application of color to displays. Proceedings of the Society for Information Display, first quarter 1979, 20, 1-2.

Harris, D. & Nelson-Heem, L. (editors). Proceedings of National Educational Computing Conference, 1981, (ED 207 526).

Heck, M. PFS: Graph. Interface Age, May 1983, 8, 86-91.

Helwig, J. SAS introductory guide. North Carolina: SAS Institute Inc., 1978.

Helwig, J. & Council, K. (editors). SAS/graph user's guide, 1981 edition. North Carolina: SAS Institute Inc., 1981.

Hillier, F. & Lieberman, G. Operations research. San Francisco: Holden-Day, Inc., 1974.

Hines, W. & Montgomery, D. Probability and statistics in engineering and management science, second edition. New York: John Wiley and Sons, 1980.

Hitt, W. An evaluation of five different abstract coding methods. Human Factors, July 1961, 3, 120-130.

Holsman, A., Hannah, L. & Schaef, E. An experimental investigation of intelligence information display parameters. RADC-TDR-63-378. Griffis Air Force Base, New York: Rome Air Development Center, September 1963.

Ives, B., Hamilton, S. & Davis, G. A framework for research in computer-based management information systems. Management Science, September 1980, 26, 910-934.

Jelden, D. The microcomputer as an interactive instruction system in the classroom: A four year staff study. Colorado: University of Northern Colorado, 1980, (ERIC 194 710).

Jessen, S. & Pannill, K. Device independent color graphics language for micros. Computer Design, July 1982, 21, 137-141.

Jones, M. Color coding. Human Factors, December 1962, 4, 355-365.

Judd, D. & Wyszecki, G. Color in business, science, and industry. New York: John Wiley & Sons, Inc., 1975.

Kanarick, A. & Petersen, R. Redundant color coding and keeping-track performance. Human Factors, April 1971, 13, 183-188.

Karner, C. Perceived versus actual value of color coding. Proceedings, Human Factors Society, 19th Annual Meeting. Santa Monica, California: Human Factors Society, 1975, 227-231.

Keister, R. Data entry performance on color versus monochromatic displays. Proceedings, Human Factors Society. Rochester, New York, October 1981, 736-740.

Kelley, N. Computer graphics: Info at a glance. Infosystems, December 1979, 26, 37-39.

Kendall, M. & Stuart, A. The advanced theory of statistics. London: Griffin & Company Limited, 1958.

King, W. & Rodriguez, J. Evaluating management information systems. Management Information Systems Quarterly, September 1978, 2, 43-51.

Kirby, P. Desktop graphics. Datamation, May 1979, 25, 164-167.

Langridge, R., Ferrin, T., Kuntz, I. & Connolly, M. Real-time color graphics in studies of molecular interactions. Science, February 13, 1981, 211, 661-666.

Leith, L. & O'Shea, F. Microcomputer graphics in education. European Association for Microprocessing and Microprogramming Journal, July 1980, 6, 205-208.

Levene, H. Robust tests for equality of variances in Contributions to probability and statistics. Olkin, I. (editor). California: Stanford University Press, 1960.

Lewis, B. & Ford, R. Basic statistics using SAS. New York: West Publishing Company, 1983.

Linksz, A. An Essay on color vision and clinical color-vision tests. New York: Grune and Stratton, 1964.

Loceff, M. Computer graphics...a picture is worth 1k words. Interface Age, February 1981, 6, 66-68.

Loceff, M. & Loceff, A. Computer graphics: Manufacturer by manufacturer. Interface Age, February 1981, 6, 78-82.

Lucas, H. Performance and the use of an information system. Management Science, April 1975, 21, 908-919.

Lucas, H. An experimental investigation of the use of computer-based graphics in decision making. Management Science, July 1981, 27, 757-768.

Lucas, H. & Nielson, N. The impact of information presentation on learning and performance. Management Science, October 1980, 26, 982-993.

Marcus, A. Images, images, images. Proceedings, Human Factors Society. Los Angeles, 1981, 219-221.

Mason, R. & Mitroff, J. A program for research in management information systems. Management Science, January 1973, 19, 475-485.

McLamb, K. Persuasion made easy with powerful presentations. Personal Computing, February 1982, 6, 35-39.

Mendenhall, W. & Scheaffer, R. Mathematical statistics with applications. North Scituate: Duxbury Press, 1973.

Miastkowski, S. & Wrege, R. Computer graphics. Popular Computing, December 1981, 1, 70-75.

Miles, E. Some highlights on low-resolution color graphics. IEEE Computer Graphics and Applications, May 1982, 2, 8-9.

Miller, F. CRT terminals get smarter, cheaper. Infosystems, September 1981, 28, 100-102.

Miller, F. The world of CRT terminals. Infosystems, June 1982, 29, 86-92.

Miller, I. Computer graphics for decision making. Harvard Business Review, November-December 1969, 47, 121-132.

Miller, J. & Rockwell, T. The use of human subjects in human factors research. Human Factors, February 1972, 35-40.

Montgomery, D. Design and analysis of experiments. New York: John Wiley and Sons, Inc., 1976.

Morris, J. Using color in industrial control graphs. Control Engineering, July 1979, 26, 41-45.

Morris, W. Decision Analysis. Columbus, Ohio: Grid, Inc., 1977.

Munns, M. Some effects of display symbol variation upon operator performance in aircraft interception. Perceptual and Motor Skills, 1968, 26, 1215-1221.

Needle, D. Lights, color, action! Graphics make computer screens come alive. Infoworld, June 21, 1982, 4, 10-11.

Neter, J. & Wasserman, W. Applied linear statistical models. Illinois: Richard Irwin, Inc., 1974.

Nickles, H. Possible effect of color on computer assisted learning. Proceedings of the National Educational Computing Conference, 1981, 191-204, (ERIC 207 526).

Park, C. Color graphics in risk analysis. Industrial Engineering Conference Proceedings, Fall 1982.

Parsons, H. The scope of human factors in computer-based data processing systems. Human Factors, April 1970, 12, 165-175.

Ramsey, H. & Atwood, M. Human factors in computer systems: A review of the literature. Catalog of Selected Documents in Psychology, May 1980, 10.

Ramsey, H., Atwood, M. & Kirschbaum, P. A critically annotated bibliography of the literature on human factors in computer systems. Technical Report SAI-78-070-DEN. Englewood, Colorado: Science Applications, Inc., May 1978.

Ray, A. (editor). SAS user's guide: Statistics, 1982 edition. North Carolina: SAS Institute Inc., 1982.

Reinhardt, P. (editor). The SAS supplemental library user's guide, 1980 edition. North Carolina: SAS Institute Inc., 1980.

Richard, G. In defense of the color computer. Popular Computing, 1982, 1, 14.

Riesenfeld, R. Current trends in computer graphics. Computers & Graphics, 1978, 3, 115-122.

Roark, D. Toward realtime interactive color graphics. Computer Design, July 1982, 21, 167-173.

Saenz, N. & Riche, C. Shape and color as dimensions of a visual redundant code. Human Factors, 1974, 16, 308-313.

Schmid, C. & Schmid, S. Handbook of graphic presentation. New York: John Wiley and Sons, Inc., 1979.

Schrack, G. Current literature in computer graphics and interactive techniques: References. Computer Graphics, February 1980, 13, 381-394.

Schrack, G. Current literature in computer graphics and interactive techniques: References, fourth series. Computer Graphics, July 1980, 14, 73-83.

Schrack, G. Computer graphics: A key word indexed bibliography for the years 1976, 1977 and 1978. Computer Graphics and Image Processing, September 1980, 14, 24-79.

Schrack, G. Computer graphics: A key word indexed bibliography for the year 1979. Computer Graphics and Image Processing, January 1981, 15, 45-78.

Schrack, G. Current literature in computer graphics and interactive techniques: References, fifth series. Computer Graphics, March 1981, 14, 200-218.

Schrack, G. Computer Graphics: A key word indexed bibliography for the year 1980. Computer Graphics and Image Processing, February 1982, 18, 145-187.

Schutz, H. An evaluation of methods for presentation of graphic multiple trends. Human Factors, July 1961, 3, 108-119.

Scott, W. Electronic displays dominating designs. Aviation Week, March 14, 1983, 118, 251-253.

Shelton, B. Influence of anxiety, aptitude, sex and cognitive development on computer-assisted learning. Dissertation Abstracts International, May 1980, vol 40 #11, 5393b.

Shontz, W., Trumm, G. & Williams, L. Color coding for information location. Human Factors, June 1971, 13, 237-246.

Skier, K. Radio Shack color computer. Popular Computing, 1982, 1, 80.

Smith, S. Color coding and visual search. Journal of Experimental Psychology, 1962, 64, 434-440.

Smith, S., Farquhar, B. & Thomas, D. Color coding in formatted displays. Journal of Applied Psychology, 1965, 49, 393-398.

Smith, S. & Thomas, D. Color versus shape coding in information displays. Journal of Applied Psychology, 1964, 48, 137-146.

Spuck, D. An analysis of the cost-effectiveness of CAI and factors associated with its successful implementation in higher education. Association for Educational Data Systems Journal, Fall 1981, 15, 10-22.

Sternecker, G. The apple graphics tablet. Creative Computing, January 1981, 7, 28-29.

Takeuchi, H & Schmidt, A. New promise of computer graphics. Harvard Business Review, January-February 1980, 58, 122-131,

Taylor, M. Color CRT terminals gain wider popularity. Mini-Micro Systems, December 1979, 12, 86-90.

Teichner, W. Color and visual information-coding. Proceedings of the Society for Information Display, 1979, 20, 3-9.

Teichner, W., Christ, R. & Corso, G. Color research for visual displays. Technical Report ONR-CR213-102-4F. Las Cruces, New Mexico: New Mexico State University, Department of Psychology, June 1977.

Teichner, W. & Krebs, M. Visual search for simple targets. Psychological Bulletin, 1974, 81, 15-28.

Truckenbroad, J. Effective use of color in computer graphics. Computer Graphics (ACM), August 1981, 15, 83-90.

Tullis, T. An evaluation of alphanumeric, graphic and color information displays. Human Factors, October 1981, 23, 541-550.

Van der Drift, K., Langerak, W., Moonen, J. & Vos, P. Final report of the computer assisted learning test project, translated by L. Stanley. Netherlands: State University of Leiden, March 1981, (ERIC 207 486).

Wagner, D. Color coding: An annotated bibliography. Technical Report NWC-TP-5922. China Lake, California: Naval Weapons Center, March 1977.

Wallace, B. Siggraph. Creative Computing, January 1981, 7, 44-49.

Waller, R., LeFrere, P. & McDonald-Ross, M. Do you need that second color? IEEE Transactions on Professional Communication, June 1982, 25, 80-85.

Watson, C. & Driver, R. The influence of computer graphics on the recall of information. Management Information Systems Quarterly, March 1983, 7, 45-53.

Wedell, J. & Alden, D. Color versus numeric coding in a keeping-track task: Performance under varying load conditions. Journal of Applied Psychology, 1973, 57, 154-159.

Zmud, R. An empirical investigation of the dimensionality of the concept of information. Decision Sciences, April 1978, 9, 187-195.

Zmud, R. Individual differences and MIS success: A review of the Empirican literature. Management Science, October 1979, 25, 966-979.

Appendix A

Background and Experience

BACKGROUND AND EXPERIENCE

1. Enter your student ID number:

2. How long have you been a university student?

   A. Less than 5 months
   B. 5-11 months
   C. 1 year (12-23 months)
   D. 2 years (24-35 months)
   E. 3 years (36-47 months)
   F. 4 years (48-59 months)
   G. 5 years (60-71 months)
   H. 6 years or more

3. What is your highest level of education now?

   A. GED certificate or high school equivalency
   B. High school graduate
   C. One or two years of college or vocational school (include Associate Degree)
   D. More than two years of college
   E. College degree (BA, BS, or equivalent)
   F. Graduate study but no graduate degree
   G. Master's degree
   H. Doctor's degree (Ph.D., M.D., LL.B., Ed.D., etc.)

4. In what college are you enrolled?

   A. College of Engineering and Applied Sciences
   B. College of Architecture
   C. College of Liberal Arts
   D. College of Business Administration
   E. College of Education
   F. College of Fine Arts
   G. College of Law
   H. College of Nursing
   I. College of Public Programs
   J. School of Social Work
   K. Other

5. What is your age in years?

6. Are you male or female?

   A. Male
   B. Female

7. Have you ever worked on a computer previous to today?

A. Yes
B. No

8. How do you feel about working with the computer?

A. Very comfortable
B. Comfortable
C. Hesitant
D. Uncomfortable
E. Very uncomfortable

9. Have you ever played video games?

A. Yes
B. No

10. How often did you usually play video games in the last six months?

A. Never play
B. 0-3 hours/week
C. 3-6 hours/week
D. 6-9 hours/week
E. 10 or more hours/week

11. Are the video games you play presented in

A. Green and black
B. Multicolor
C. Black and white
D. Other
E. I have never played video games

12. How often did you usually interact with the computer using a CRT screen display for non-video game activities in the last six months?

A. Never
B. 0-3 hours/week
C. 3-6 hours/week
D. 6-9 hours/week
E. 10 or more hours/week

Appendix B

Case Study

PROBLEM

Facts

You are the Sales Manager for the U-Sell-Um Bike Shop. Your store is located near the University. It is time to order your bike stock for the summer season. It is your job to determine what type of bike and how many you should buy in order to maximize profits for the U-Sell-Um Bike Shop. There are several facts that may affect your final decisions.

1. There are three models of bikes on the market:

   1. Campus Bikes
   2. 3-Speed Bikes
   3. Deluxe Bikes

2. The selling price, purchase, storage, understock (shortage) and overstock (overage) costs are known for each model.

3. Bikes can be purchased from their respective manufacturers only in multiples of 20 bikes.

4. Because of advertising, literature, sales expertise and special bike rack displays, only one model of bike can be stocked in your store during a season.

5. Presently you have room to display a maximum of 30 bikes. A friend has a warehouse that has room to store

up to 10 more bikes free of charge. However if you have more than 40 bikes in your inventory, you will have to rent space for all bikes you do not have room for in your shop display. The rental fee is $5/bike/season no matter how long it is stored.

6. Unless the environment changes (customers or competition) the demand for all styles of bike for any season is 33.

7. Presently your shop is sold out of bikes and does not have any unfilled orders.

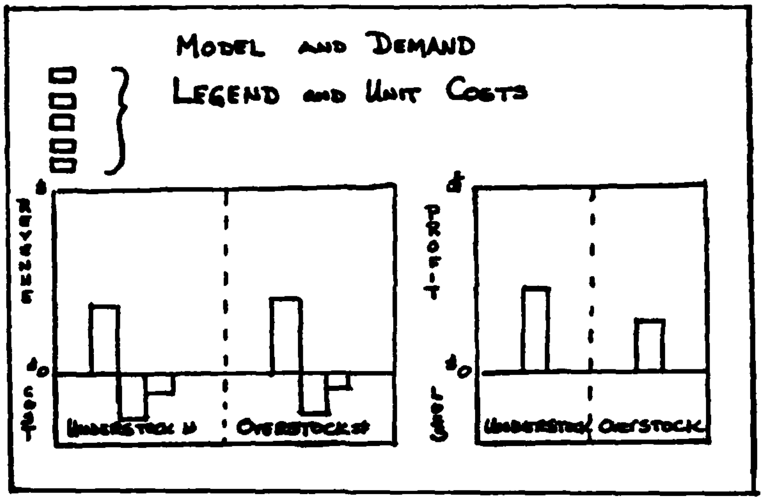
As a business person, your goal is to maximize your profits when considering the above information.

Go on to the next page.

Instructions

Revenue and cost information concerning the three bike models is given on computer generated graphics sheets located in your packet. This information is in bargraph format. Each model's figures are shown on a separate screen. You have the option of repeatedly reviewing any of the models' bargraphs during each case study.

The format on the graphics sheets for every model's information is the same. The title indicates the season, model and demand. Beneath the title is the legend which gives the per bike figures and the key to reading the charts. Under the legend are two charts. The bargraph chart on the left represents the breakdown of revenues and costs when buying two different quantities of bikes (in multiples of 20). The first set of bars represents stocking fewer than the demand (understocking) while the second set represents stocking more than the demand (overstocking). The right bargraph chart sums the left chart's revenues and costs into total profit or loss figures for each of the two stocking quantities. (See figure below.)



There are 3 different cases in which you are asked to review a group of bargraph charts and make some decisions. These decisions are in question format. All questions and answers are based on the information given on the graphics sheets and problem scenario.. The questions are printed in a multiple choice format and inserted in the appropriate case study packet. You are urged to refer to the presentations while making your decisions. Any of the presentations associated with the particular case can be viewed at any time during that particular case study. Look over all the questions before you begin your analysis. Note, once you have made a decision (answered a question), it can not be changed.

After you have finished the 3 cases you will be asked a few additional questions by the U-Sell-Um Bike Shop

regional manager, Ms. Bucks. At that time you will not have the graphical information available for review.

If you have any questions, please ask the Experimenter now. After any questions have been answered, review the Problem Facts and wait for the Experimenter to tell you to begin.

CASE 1: SUMMER POLICY PURCHASE DECISIONS

1. Based on the current sales information, which type of bike model will you stock to maximize your profits?

   A. Campus Bike Model
   B. 3-Speed Bike Model
   C. Deluxe Bike Model

2. How many of that type would you order for the summer?

   A. 20
   B. 40

3. If for some reason you stocked Campus Bikes, what is your estimate of your total profits?

   A. $210
   B. $340
   C. $500
   D. $825

4. Which of the following pairs represents the minimum and maximum dollar figures associated with overstocking?

   A. (200, 600)
   B. (90, 520)
   C. (70, 560)
   D. (150, 560)
   E. (200,1200)

5. Based on the graphical information which rule of thumb best describes the decision on the purchase of any of the bike models?

   A. Purchase less then the demand when shortage costs are less than overstocking costs.
   B. Smaller costs guarantee more profits.
   C. More inventory is always more profitable.
   D. None of the above statements are appropriate.

6. Warehouse cost associated with stocking the 3-Speed Bikes is a maximum of

   A. $ 70.
   B. $ 260.
   C. $ 600.
   D. $1200.
   E. None as the warehouse is not used

7. If you could cut costs somewhere by 10% despite the model of bike, what cost would be most profitable to cut?

   A. Purchase Cost
   B. Shortage Cost
   C. Overstock Cost
   D. Cutting any of the above costs by 10% would have the same effect.

8. If advertising would increase the demand from 33 to 37 Deluxe Bikes, what is the MAXIMUM you would be willing to spend on advertising?

   A. None as advertising would never be profitable
   B. $300
   C. $400
   D. $500
   E. $600

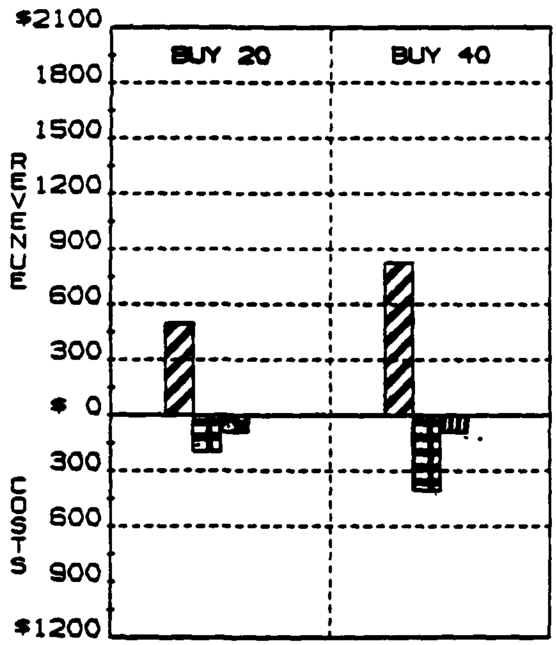
9. If you understock the Deluxe Bike, what are your expected shortage costs?

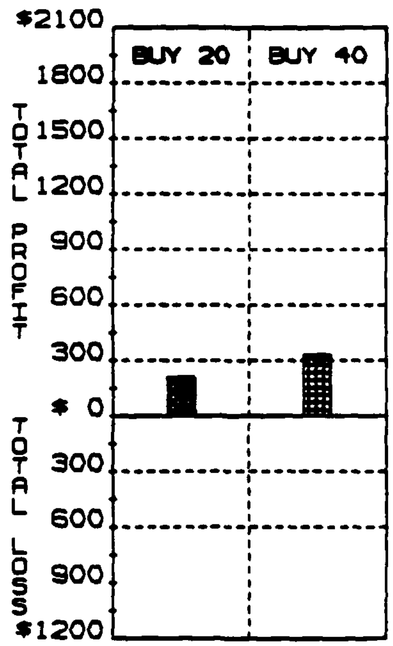
   A. $ 520
   B. $ 600
   C. $1200
   D. None as no shortages are expected

You are finished with this analysis. Please raise your hand to indicate to the Experimenter you are ready for the next case study.

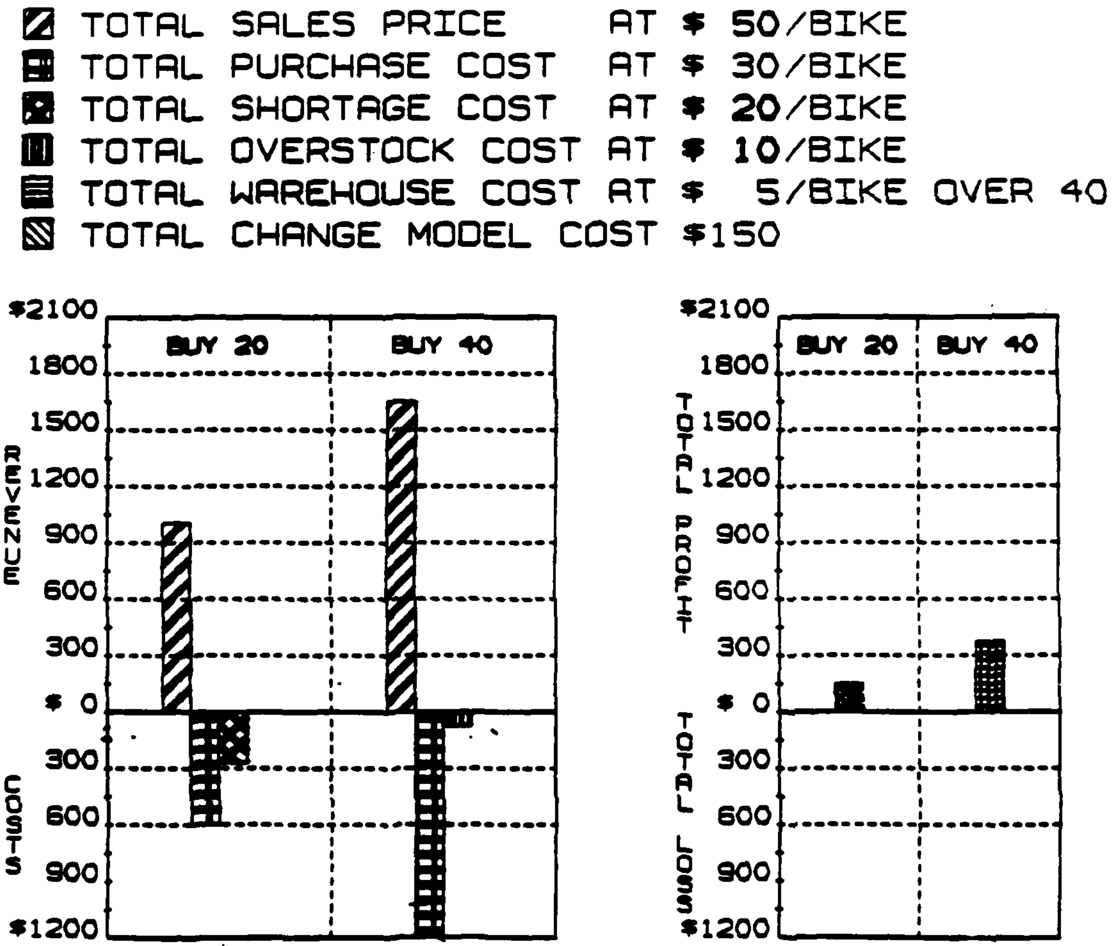
SUMMER: DEMAND 33 CAMPUS BIKES

TOTAL SALES PRICE AT $ 25/BIKE
TOTAL PURCHASE COST AT $ 10/BIKE
TOTAL SHORTAGE COST AT $ 7/BIKE
TOTAL OVERSTOCK COST AT $ 12/BIKE
TOTAL WAREHOUSE COST AT $ 5/BIKE OVER 40
TOTAL CHANGE MODEL COST $150



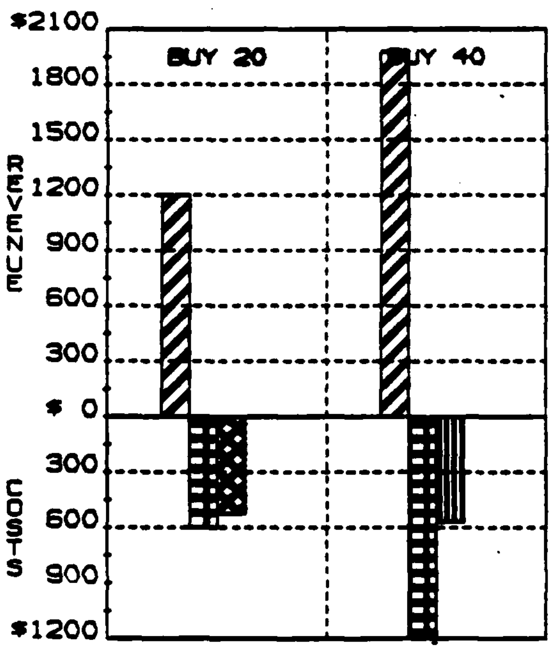


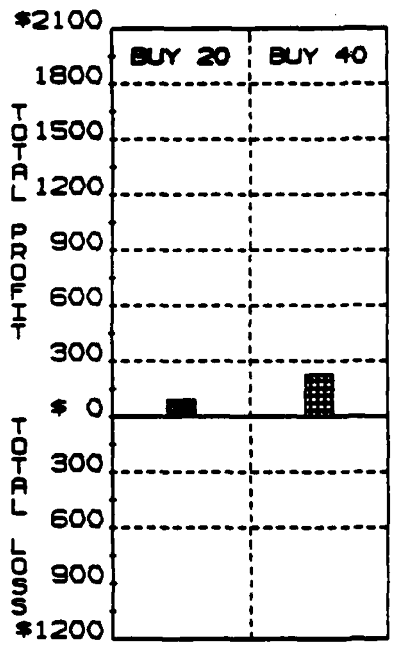
SUMMER: DEMAND 33 3-SPEED BIKES



SUMMER: DEMAND 33 DELUXE BIKES

TOTAL SALES PRICE AT $ 60/BIKE
TOTAL PURCHASE COST AT $ 30/BIKE
TOTAL SHORTAGE COST AT $ 40/BIKE
TOTAL OVERSTOCK COST AT $ 80/BIKE
TOTAL WAREHOUSE COST AT $ 5/BIKE OVER 40
TOTAL CHANGE MODEL COST $150





CASE 2 SCENARIO: FALL INVENTORY

Your manufacturers' representatives have contacted you and are requiring you to order now for both your Summer shipment and Fall shipment. The Fall demand will be affected by the University's new parking policy. Parking fees will be higher and the close-in parking lots will be reserved for faculty, staff and visitors only. The following facts should be considered when determining the Fall purchase order for bikes.

1. The demand for the Fall will change from that of the summer due to the University's new parking policy. University students prefer the cheaper Campus bike.

	Summer Demand	Fall Demand
Campus Bike	33	67
3-Speed Bike	33	27
Deluxe Bike	33	25

2. Your shop can store 30 bikes. Remember a friend will store up to 10 more for you in his warehouse for free. However if you have more than 40 in your inventory, you will have to rent space for ALL bikes you do not have room for in your shop display. The rental fee is

$5/bike/season no matter how long it is stored.

Initially you are interested only in the Fall figures that reflect the change in demand due to the University's new parking policy. The Summer inventory figures are not of concern. The revenue/cost and profit bargraphs have been updated to reflect the Fall demand options. The three presentations show the Fall options without considering the impact of the Summer purchase. Using these graphs answer Case 2 Purchase Decision Questions. Review the questions before you start the analysis. Remember all questions are based on the information given on the graphics sheets.

CASE 2: FALL POLICY PURCHASE DECISIONS

SUMMER INVENTORY NOT CONSIDERED

1. What inventory purchase would maximize your profits for the Fall?

   A. Campus/60
   B. Campus/80
   C. 3-Speed/20
   D. 3-Speed/40
   E. Deluxe/20
   F. Deluxe/40

2. Did the change of demand of the different models due to the University's new parking policy affect the model of bike you will stock in the Fall?

   A. Yes, the change in demand changes the model I will stock
   B. No, the change in demand makes the model I planned for the Summer even more profitable

3. What is the MAXIMUM you would pay to expand your display area if the future demand remains the same as in the Fall?

   A. $ 0. No extra display space is needed
   B. $150.
   C. $250.

4. How do the Fall profits compare to the Summer profits? (Answer each question T for True or F for False.)

   T F Understocking remains the rule of thumb as it is more profitable than overstocking for all models.

   T F The minimum inventory purchase considered during both seasons was 20 bikes which always guaranteed profits.

   T F Warehouse costs do not affect the total profits of the most profitable model for either season.

   T F If your inventory policy was chosen by Lady Luck you have a chance of going in debt during each season.

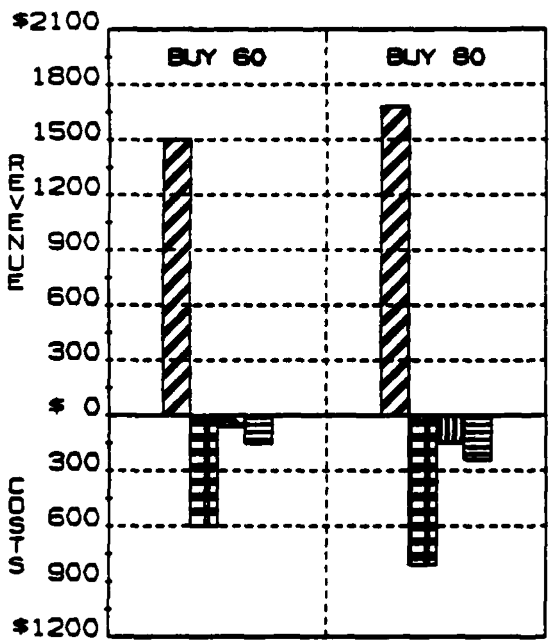
   T F It would be worth spending $100/season on advertising if it guaranteed the depletion of your stock.

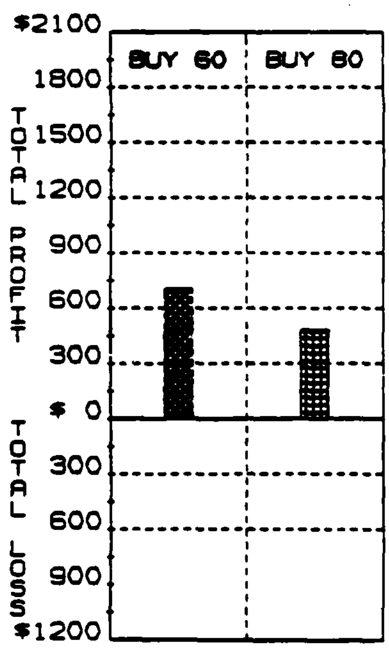
You are finished with this analysis. Please go on to the next case study.

FALL FIGURES: DEMAND 67 CAMPUS BIKES

SUMMER INVENTORY NOT CONSIDERED

- TOTAL SALES PRICE AT $ 25/BIKE
- TOTAL PURCHASE COST AT $ 10/BIKE
- TOTAL SHORTAGE COST AT $ 7/BIKE
- TOTAL OVERSTOCK COST AT $ 12/BIKE
- TOTAL WAREHOUSE COST AT $ 5/BIKE OVER 40
- TOTAL CHANGE MODEL COST $150

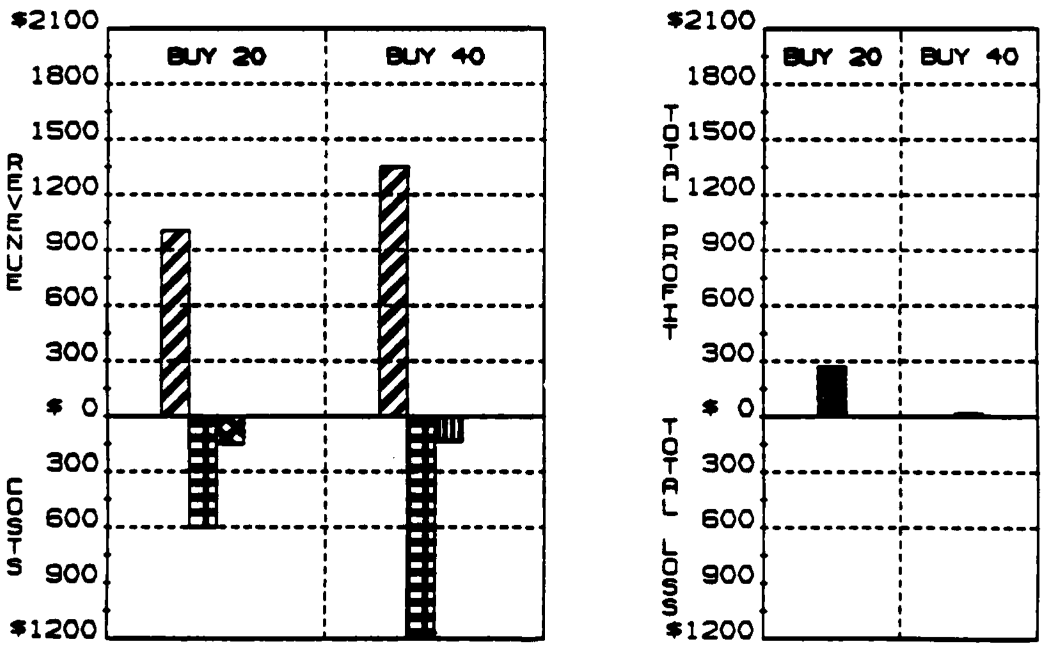




FALL FIGURES: DEMAND 27 3-SPEED BIKES

SUMMER INVENTORY NOT CONSIDERED

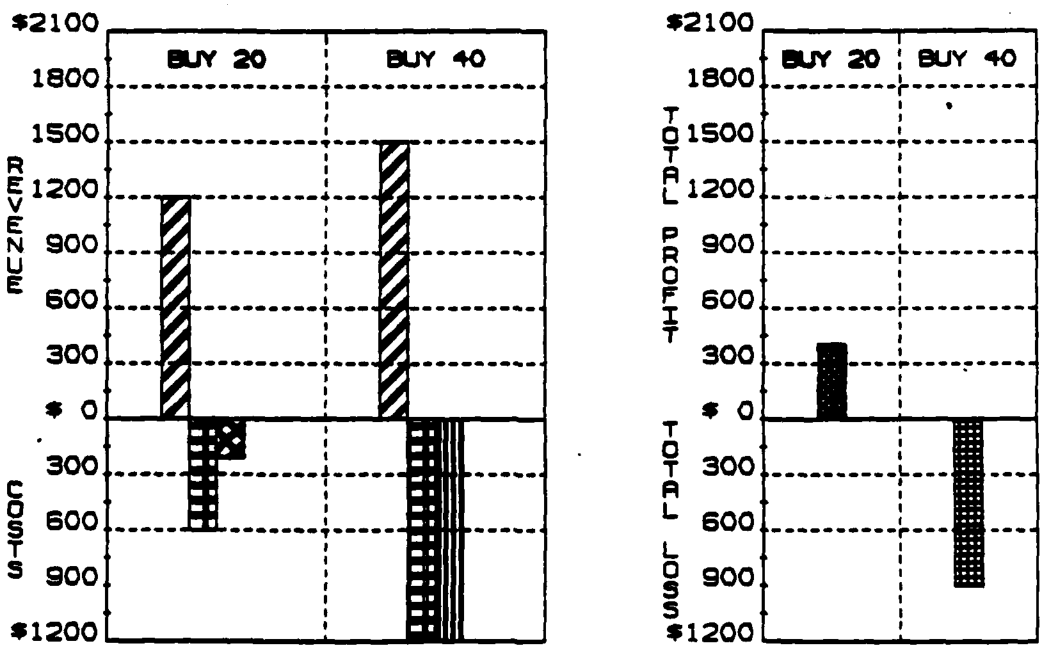
- TOTAL SALES PRICE AT $ 50/BIKE
- TOTAL PURCHASE COST AT $ 30/BIKE
- TOTAL SHORTAGE COST AT $ 20/BIKE
- TOTAL OVERSTOCK COST AT $ 10/BIKE
- TOTAL WAREHOUSE COST AT $ 5/BIKE OVER 40
- TOTAL CHANGE MODEL COST $150



FALL FIGURES: DEMAND 25 DELUXE BIKES

SUMMER INVENTORY NOT CONSIDERED

- TOTAL SALES PRICE AT $ 60/BIKE
- TOTAL PURCHASE COST AT $ 30/BIKE
- TOTAL SHORTAGE COST AT $ 40/BIKE
- TOTAL OVERSTOCK COST AT $ 80/BIKE
- TOTAL WAREHOUSE COST AT $ 5/BIKE OVER 40
- TOTAL CHANGE MODEL COST $150



CASE 3 SCENARIO: LONG RANGE PLANNING

As a practical business person you realize that the summer's inventory and sales may have a significant impact on your Fall purchasing decisions. Such things as summer shortages, overstocking, warehouse costs and costs of changing models of bikes may have an effect on what bike model becomes most profitable to stock in the Fall. After looking into these particular costs, you have found out the following facts:

1. Any shortages of a particular stocked model from the Summer will add to the Fall demand if the model is not changed between seasons.

2. Any overstocking of a particular model from the Summer will add to the Fall inventory if the model is not changed between seasons.

3. If the model is changed between seasons, any shortages or overstocking will not affect the Fall figures. Shortages are lost business and overstocked bikes will be donated to local charity.

4. Changing models between seasons costs $150. This expense covers the change in customer literature, advertising, display racks, etc.

5. A friend's warehouse is still available for

storing 10 bikes free. However if the total inventory is over 40, you have to contract for storage space for all bikes that can not be displayed in your store. The storage fee is still $5/bike.

The revenue/costs and profits charts have been updated to reflect the impact any summer inventory purchases have on the Fall figures. Using these graphs, answer Case 3 Purchasing Decision questions.

CASE 3: LONG RANGE PLANNING

IMPACT OF SUMMER INVENTORY ON FALL PURCHASES

1. When considering the Summer purchase inventory, does it appear profitable to change models from Summer to Fall?

   A. Yes
   B. No

2. When considering the impact of the Summer inventory, what inventory purchase would maximize your profits in the Fall?

   A. Campus/60
   B. Campus/80
   C. 3-Speed/20
   D. 3-Speed/40
   E. Deluxe/0
   F. Deluxe/20
   G. Deluxe/40

3. When considering the Summer inventory, does the University's new policy have an effect on your Fall purchase decision?

   A. Yes
   B. No

4. When considering the Summer inventory in your Fall decision, how much is your friend's warehouse offer worth to you in dollars saved?

   A. $ 0 His warehouse is never used
   B. $15
   C. $50
   D. over $65

5. The Campus model manufacturer's representative promises that her company will pay all change-over expenses ($150) necessary to insure your shop will stock Campus bikes in the Fall. Will her offer affect your purchase decision?

   A. No Campus Bikes was the model originally planned for purchase
   B. No Campus Bikes still would not be the most profitable to stock
   C. Yes Eliminating the change-over cost for Campus Bikes makes that model most profitable

6. Your regional manager, Ms. Bucks, just called and wants you to stock the Deluxe model for both the Summer and Fall. Would it be profitable for you to invest $50 in advertising to insure all the Deluxe bikes are sold by the end of the Fall season?

   A. No The most profitable inventory purchase has shortages so advertising would be an added expense
   B. No The advertising cost would NOT save over $50 in overstocking costs
   C. Yes The advertising cost would save over $50 in overstocking costs

7. When considering the Summer inventory are you guaranteed to make a profit no matter what Fall purchase policy you chose?

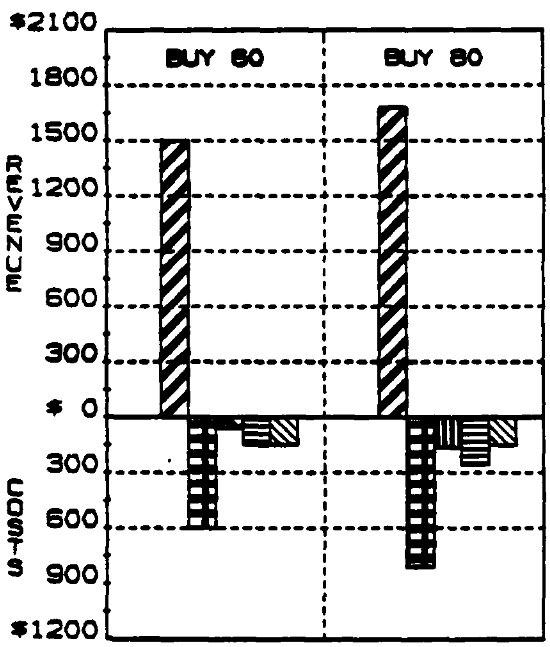
   A. Yes
   B. No

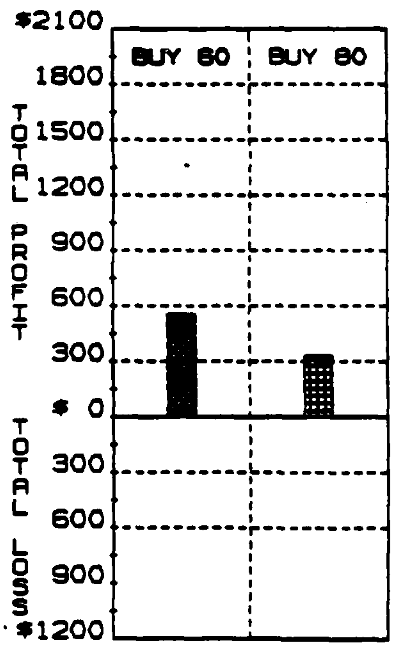
You are finished with the interactive decision making session. Please raise your hand to indicate to the Experimenter you are ready for the final questionnaire.

FALL FIGURES: DEMAND 67 CAMPUS BIKES

SUMMER PURCHASE NO CAMPUS BIKES

TOTAL SALES PRICE AT $ 25/BIKE
TOTAL PURCHASE COST AT $ 10/BIKE
TOTAL SHORTAGE COST AT $ 7/BIKE
TOTAL OVERSTOCK COST AT $ 12/BIKE
TOTAL WAREHOUSE COST AT $ 5/BIKE OVER 40
TOTAL CHANGE MODEL COST $150

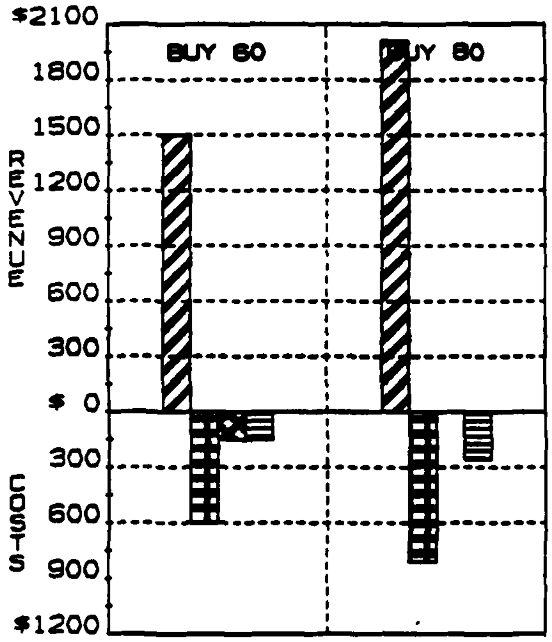


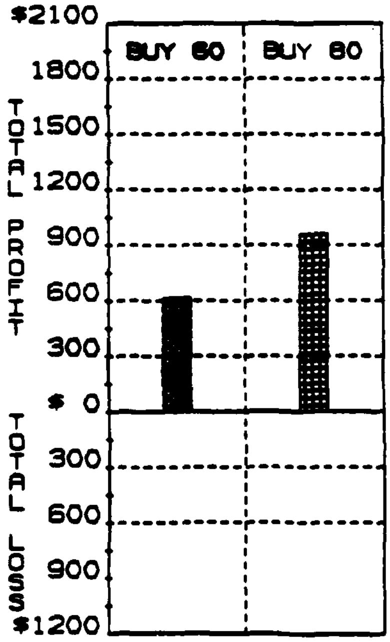


FALL FIGURES: DEMAND 67 CAMPUS BIKES

SUMMER PURCHASE 20 CAMPUS BIKES

- TOTAL SALES PRICE AT $ 25/BIKE
- TOTAL PURCHASE COST AT $ 10/BIKE
- TOTAL SHORTAGE COST AT $ 7/BIKE
- TOTAL OVERSTOCK COST AT $ 12/BIKE
- TOTAL WAREHOUSE COST AT $ 5/BIKE OVER 40
- TOTAL CHANGE MODEL COST $150

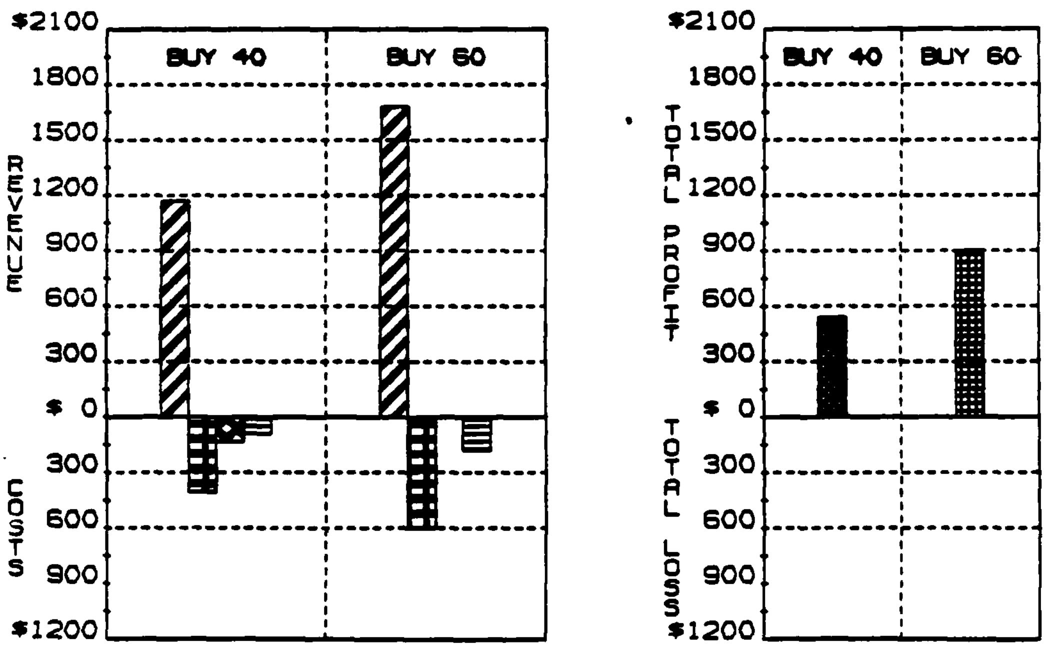




FALL FIGURES: DEMAND 67 CAMPUS BIKES

SUMMER PURCHASE 40 CAMPUS BIKES

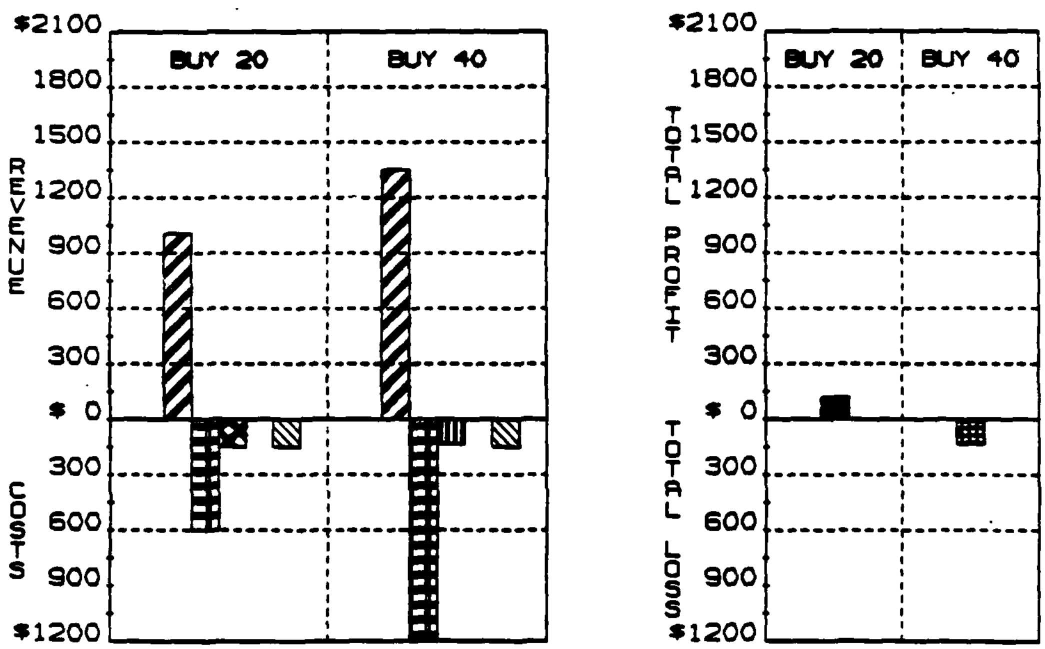
- TOTAL SALES PRICE AT $ 25/BIKE
- TOTAL PURCHASE COST AT $ 10/BIKE
- TOTAL SHORTAGE COST AT $ 7/BIKE
- TOTAL OVERSTOCK COST AT $ 12/BIKE
- TOTAL WAREHOUSE COST AT $ 5/BIKE OVER 40
- TOTAL CHANGE MODEL COST $150



FALL FIGURES: DEMAND 27 3-SPEED BIKES

SUMMER PURCHASE NO 3-SPEED BIKES

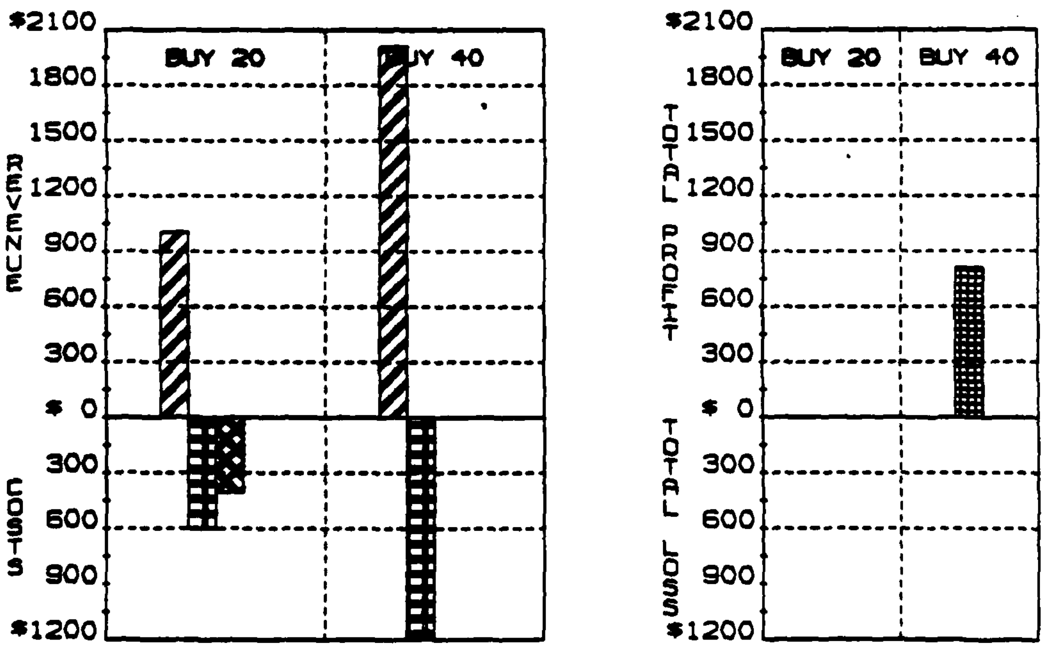
TOTAL SALES PRICE AT $ 50/BIKE
TOTAL PURCHASE COST AT $ 30/BIKE
TOTAL SHORTAGE COST AT $ 20/BIKE
TOTAL OVERSTOCK COST AT $ 10/BIKE
TOTAL WAREHOUSE COST AT $ 5/BIKE OVER 40
TOTAL CHANGE MODEL COST $150



FALL FIGURES: DEMAND 27 3-SPEED BIKES

SUMMER PURCHASE 20 3-SPEED BIKES

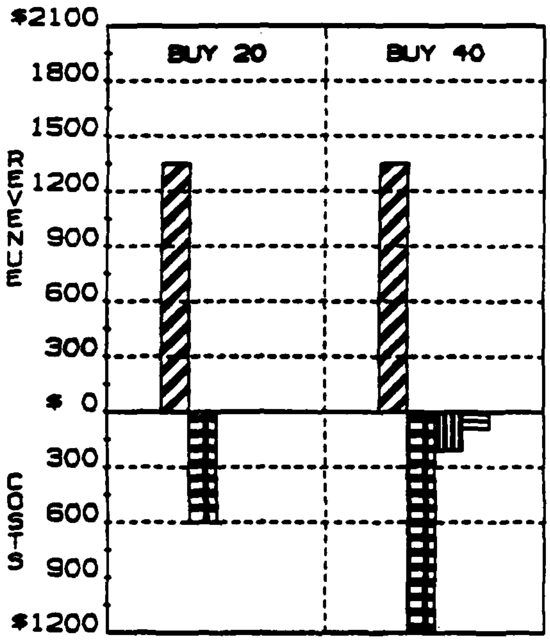
- TOTAL SALES PRICE AT $ 50/BIKE
- TOTAL PURCHASE COST AT $ 30/BIKE
- TOTAL SHORTAGE COST AT $ 20/BIKE
- TOTAL OVERSTOCK COST AT $ 10/BIKE
- TOTAL WAREHOUSE COST AT $ 5/BIKE OVER 40
- TOTAL CHANGE MODEL COST $150

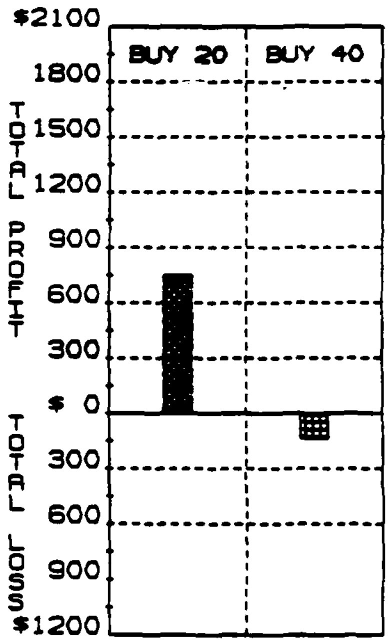


FALL FIGURES: DEMAND 27 3-SPEED BIKES

SUMMER PURCHASE 40 3-SPEED BIKES

TOTAL SALES PRICE AT $ 50/BIKE
TOTAL PURCHASE COST AT $ 30/BIKE
TOTAL SHORTAGE COST AT $ 20/BIKE
TOTAL OVERSTOCK COST AT $ 10/BIKE
TOTAL WAREHOUSE COST AT $ 5/BIKE OVER 40
TOTAL CHANGE MODEL COST $150

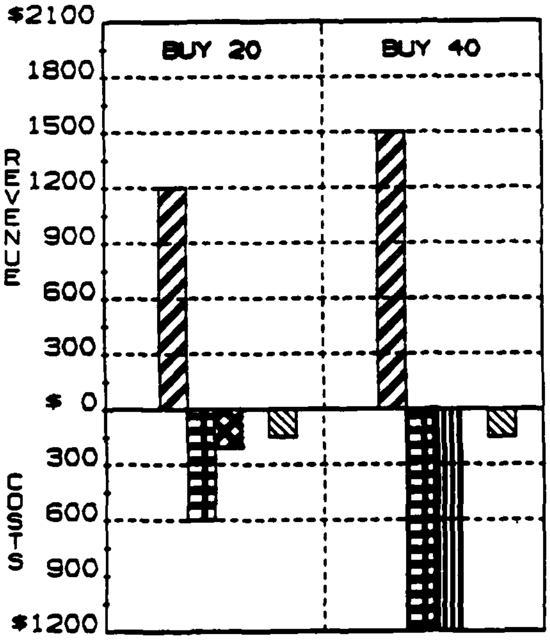


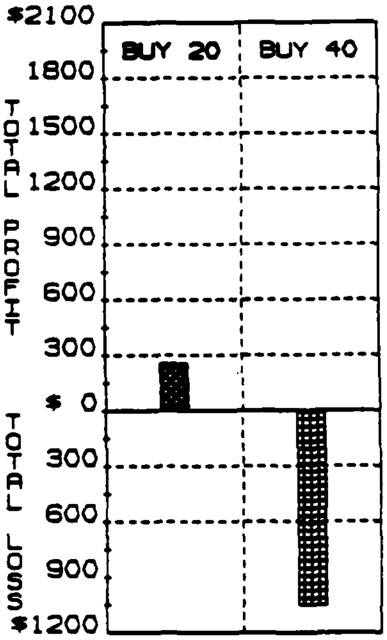


FALL FIGURES: DEMAND 25 DELUXE BIKES

SUMMER PURCHASE NO DELUXE BIKES

- TOTAL SALES PRICE AT $ 60/BIKE
- TOTAL PURCHASE COST AT $ 30/BIKE
- TOTAL SHORTAGE COST AT $ 40/BIKE
- TOTAL OVERSTOCK COST AT $ 80/BIKE
- TOTAL WAREHOUSE COST AT $ 5/BIKE OVER 40
- TOTAL CHANGE MODEL COST $150

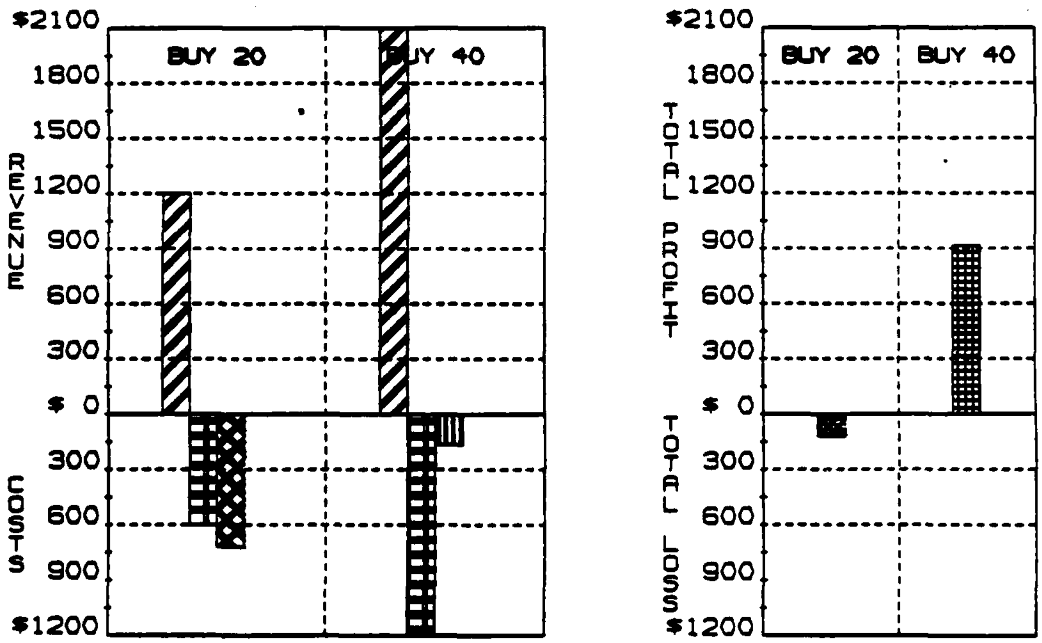




FALL FIGURES: DEMAND 25 DELUXE BIKES

SUMMER PURCHASE 20 DELUXE BIKES

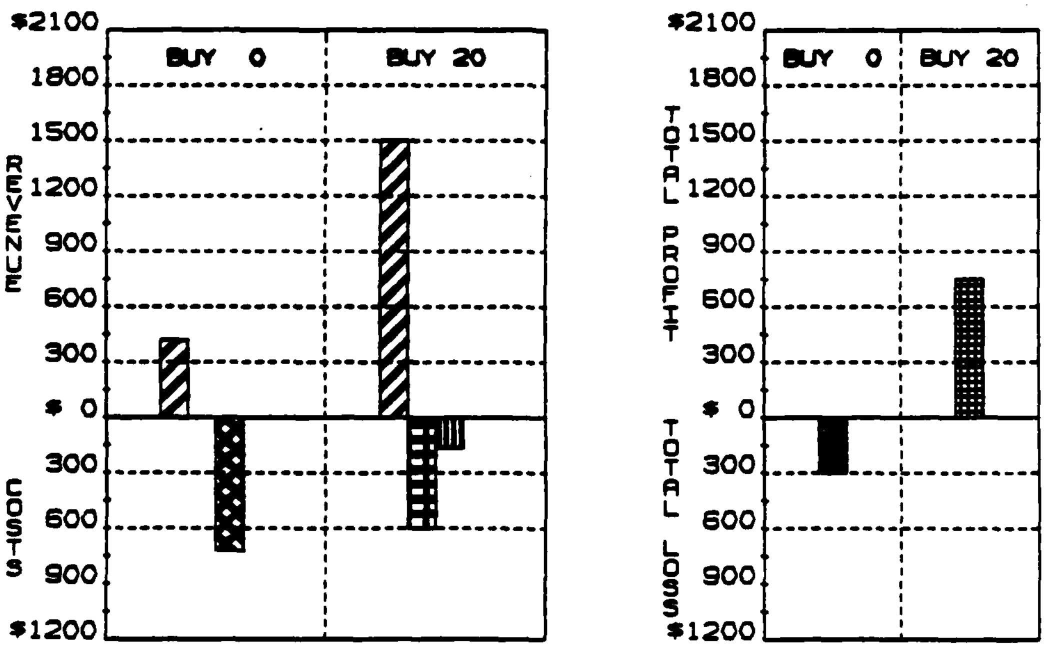
- TOTAL SALES PRICE AT \$ 60/BIKE
- TOTAL PURCHASE COST AT \$ 30/BIKE
- TOTAL SHORTAGE COST AT \$ 40/BIKE
- TOTAL OVERSTOCK COST AT \$ 80/BIKE
- TOTAL WAREHOUSE COST AT \$ 5/BIKE OVER 40
- TOTAL CHANGE MODEL COST \$150



FALL FIGURES: DEMAND 25 DELUXE BIKES

SUMMER PURCHASE 40 DELUXE BIKES

- TOTAL SALES PRICE AT $ 60/BIKE
- TOTAL PURCHASE COST AT $ 30/BIKE
- TOTAL SHORTAGE COST AT $ 40/BIKE
- TOTAL OVERSTOCK COST AT $ 80/BIKE
- TOTAL WAREHOUSE COST AT $ 5/BIKE OVER 40
- TOTAL CHANGE MODEL COST $150



MEETING WITH MS. BUCKS

The U-Sell-Um regional manager, Ms. Bucks, has stopped by the shop to see how business is going. She is interested in your future plans. She has contacted the manufacturers and has some information concerning possible contracts with them. She wishes to discuss the information with you before you send in your purchase order.

1. When only the summer inventory was considered, what model appeared the most profitable to stock for the Summer?

   A. Campus bikes
   B. 3-Speed Bikes
   C. Deluxe Bikes

2. When taking into account only the summer inventory and demand, what model appeared to be the most profitable to stock for the Fall?

   A. Campus Bikes
   B. 3-Speed Bikes
   C. Deluxe Bikes

3. Ms. Bucks reminds you that your ultimate goal is to maximize total profits over both seasons. Which of the following describes your purchase decisions for the next two seasons? (Circle T for True, F for False)

   T F Maximizing each season's profits maximizes total profits.

   T F Warehouse costs are the significant cost factor in determing the most profitable model(s) to stock for the next two seasons.

   T F Warehouse costs are the significant cost factor in determining the most profitable quantities to stock for the next two seasons.

   T F When total supply of bikes is the same as total demand, maximum profits are guaranteed over the next two seasons if overstocking is the summer's policy.

4. How much profit do you expect to make over the next two seasons?

   A. Approximately $ 800
   B. Approximately $1000
   C. Approximately $1200
   D. Approximately $1400

5. What model of bike do you plan to purchase for your Summer inventory in order to maximize your total profits?

   A. Campus Bikes
   B. 3-Speed Bikes
   C. Deluxe Bikes

6. What model of bike do you plan to purchase for your Fall inventory in order to maximize your total profits?

   A. Campus Bikes
   B. 3-Speed Bikes
   C. Deluxe Bikes

7. Is your friend's warehouse offer worth enough (over $0) to you to mention it to Ms. Bucks?

   A. Yes
   B. No

8. What was the overage cost/bike associated with the 3-Speed model?

   A. $ 7/bike
   B. $10/bike
   C. $20/bike
   D. $50/bike
   E. $80/bike

9. What was the shortage cost/bike associated with the Deluxe model?

   A. $ 5/bike
   B. $10/bike
   C. $20/bike
   D. $40/bike
   E. $80/bike

10. The dollar values printed on the scale of the bargraphs in each chart were in multiples of

   A. $100
   B. $200
   C. $300
   D. $500

11. Ms. Bucks forecasts a new line of bike coming late Summer or early Fall. Because she is uncertain when they will become available on the open market, she wants to be sure you have room for them as she feels they will be the hottest bike ever. Therefore she insists you understock both seasons. If you do, will you be guaranteed profits no matter what model(s) you stock?

    A. Yes Understocking is the rule of thumb for the next two seasons
    B. Yes Understocking maximizes profits as long as models are not changed between seasons
    C. No Shortages can turn profits into losses even when the same model is stocked during both seasons

Ms. Bucks has to leave for another store. She thanks you for your information and wishes you the best of luck over the next two seasons.

You are finished with this experiment. Thank you for your generous cooperation.

Appendix C

Fortran/DI 3000 Program

```
      PROGRAM DISSER
C****PROGRAM DISSER INITIALIZES SYSTEM WITH DISK.  IT THEN GIVES CONTROL
C            TO SUBROUTINE GROUP
      INTEGER IX, NPRES, K
      COMMON/INPUT/IX, NPRES, K
C
C*******GIVE CONTROL TO APPROPRIATE GROUP
      K=1
      IX=1
1     CONTINUE
      IF (IX.GT.3) GO TO 2
C*******NOT FINISHED WITH GROUPS.  CALL NEXT GROUP
      CALL GROUP
      WRITE (3,'(1X,"   ")')
      GO TO 1
C*******ALL GROUPS PROCESSED.  DISPLAY FINISH MSG TO SUBJECT
2     CONTINUE
100   WRITE (3,'(1X"   YOU ARE FINISHED USING THE CRT SCREEN")')
      WRITE (3,'(1X,"    IT IS TIME TO REPORT TO MS. LUCAS")')
      WRITE (3,'(1X,"    PLEASE COMPLETE THE LAST QUESTIONNAIRE")')
     WRITE (3,'(1X,"   ")')
     WRITE (3,'(1X,"   ")')
      WRITE (3,'(1X,"   "I2)')
C*******TERMINATE PROGRAM
      STOP
      END
C
C
C
C
      SUBROUTINE GROUP
      COMMON/INPUT/IX, NPRES, K
C
C*******INSURE IX IS NOT GREATER THAN 3
10    GO TO (100,300,700),IX
C*******GROUP NUMBER IS OUT OF RANGE
20    RETURN
C
C
C*****VIEWGRAPH OPTIONS
100   WRITE (3,'(1X,"   ")')

14    IF (K.NE.1) GO TO 150
C*******RESET START TIME FOR FIRST CASE
      CALL STIME
150   WRITE (3,'(1X,"   ENTER VIEWGRAPH NUMBER YOU WANT TO SEE")')
      WRITE (3,'(1X,"    THEN HIT THE RETURN KEY")')
      WRITE (3,'(1X,"
     ")')
      WRITE (3,'(1X,"   1   SUMMER DEMAND 35 CAMPUS BIKES")')
      WRITE (3,'(1X,"   2   SUMMER DEMAND 33 3-SPEED BIKES")')
      WRITE (3,'(1X,"   3   SUMMER DEMAND 33 DELUXE BIKES")')
      WRITE (3,'(1X,"   10, NONE, I AM FINISHED WITH THIS ANALYSIS")'
     ),
      WRITE (3,'(1X,"   ENTER 1, 2, 3 OR 10")')
      WRITE (3,'(1X,"    THEN HIT THE RETURN KEY")')
      READ (3,*) NPRES
      IF (NPRES.EQ.1) GO TO 400
      IF (NPRES.EQ.2) GO TO 430
      IF (NPRES.EQ.3) GO TO 430
      IF (NPRES.EQ.10) GO TO 400
C*******SUBJECT DID NOT FOLLOW DIRECTIONS.  PROMPT AGAIN CLEANING SCREEN
160   CONTINUE
      WRITE (3,'(1X," ")')
      WRITE (3,'(1X,"   THE NUMBER YOU ENTERED IS NOT AVAILABLE")')
      GO TO 150
C
C
C*****GROUP 2 OPTIONS
300   WRITE (3,'(1X," ")')
350   WRITE (3,'(1X,"   ENTER THE VIEWGRAPH NUMBER YOU WANT TO SEE")')
      WRITE (3,'(1X,"
     ")')
      WRITE (3,'(1X,"   1   FALL DEMAND 27 CAMPUS BIKES")')
      WRITE (3,'(1X,"        SUMMER INVENTORY NOT CONSIDERED")')
      WRITE (3,'(1X,"
     ")')
      WRITE (3,'(1X,"   2   FALL DEMAND 27 3-SPEED BIKES")')
      WRITE (3,'(1X,"        SUMMER INVENTORY NOT CONSIDERED")')
      WRITE (3,'(1X,"
     ")')
```

```
          WRITE (3,'(1X,"   3  FALL DEMAND 25 DELUXE BIKES")')
          WRITE (3,'(1X,"      SUMMER INVENTORY NOT CONSIDERED")')
          WRITE (3,'(1X," ")')
           WRITE (3,'(1X,"  10  NONE, I AM FINISHED WITH THIS ANALYSIS")')
          WRITE (3,'(1X," ")')
          WRITE (3,'(1X,"  ENTER 1, 2, 3 OR 10")')
           WRITE (3,'(1X,"   THEN HIT THE RETURN KEY")')
          READ (3,*) NPRES
          IF (NPRES.EQ.1) GO TO 400
          IF (NPRES.EQ.2) GO TO 400
          IF (NPRES.EQ.3) GO TO 400
          IF (NPRES.EQ.10) GO TO 400
C*****SUBJECT DID NOT FOLLOW DIRECTIONS.  PROMPT AGAIN CLEARING SCREEN
        CONTINUE
          WRITE (3,'(1X," ")')
           WRITE (3,'(1X,"   THE NUMBER YOU ENTERED IS NOT AVAILABLE")')
          GO TO 300
C
C
C*****GROUP 3 OPTIONS
700       WRITE (3,'(1X," ")')
          WRITE (3,'(1X,"   ENTER THE VIEWGRAPH NUMBER YOU WANT TO SEE")')
            WRITE (3,'(1X," ")')
            WRITE (3,'(1X,"   1  FALL DEMAND 67 CAMPUS BIKES")')
            WRITE (3,'(1X,"      SUMMER PURCHASE 20 CAMPUS BIKES")')
            WRITE (3,'(1X," ")')
            WRITE (3,'(1X,"   2  FALL DEMAND 67 CAMPUS BIKES")')
            WRITE (3,'(1X,"      SUMMER PURCHASE 40 CAMPUS BIKES")')
            WRITE (3,'(1X," ")')
            WRITE (3,'(1X,"   3  FALL DEMAND 67 CAMPUS BIKES")')
            WRITE (3,'(1X,"      SUMMER PURCHASE NO CAMPUS BIKES")')
            WRITE (3,'(1X," ")')
            WRITE (3,'(1X,"   4  FALL DEMAND 27 3-SPEED BIKES")')
            WRITE (3,'(1X,"      SUMMER PURCHASE 20 3-SPEED BIKES")')
            WRITE (3,'(1X," ")')
            WRITE (3,'(1X,"   5  FALL DEMAND 27 3-SPEED BIKES")')
            WRITE (3,'(1X,"      SUMMER PURCHASE 40 3-SPEED BIKES")')
            WRITE (3,'(1X," ")')
            WRITE (3,'(1X,"   6  FALL DEMAND 27 3-SPEED BIKES")')
            WRITE (3,'(1X,"      SUMMER PURCHASE NO 3-SPEED BIKES")')
            WRITE (3,'(1X," ")')
            WRITE (3,'(1X,"   7  FALL DEMAND 25 DELUXE BIKES")')
            WRITE (3,'(1X,"      SUMMER PURCHASE 20 DELUXE BIKES")')
            WRITE (3,'(1X," ")')
            WRITE (3,'(1X,"   8  FALL DEMAND 25 DELUXE BIKES")')
            WRITE (3,'(1X,"      SUMMER PURCHASE 40 DELUXE MODELS")')
            WRITE (3,'(1X," ")')
            WRITE (3,'(1X,"   9  FALL DEMAND 25 DELUXE BIKES")')
            WRITE (3,'(1X,"      SUMMER PURCHASE NO DELUXE BIKES")')
            WRITE (3,'(1X," ")')
            WRITE (3,'(1X,"  10  NONE, I AM FINISHED WITH THIS ANALYSIS")')
            WRITE (3,'(1X," ")')
            WRITE (3,'(1X,"  ENTER 1, 2, 3, 4, 5, 6, 7,")')
             WRITE (3,'(10X,"8,  9  OR 10")')
             WRITE (3,'(1X,"   THEN HIT THE RETURN KEY")')
          READ (3,*) NPRES
          IF (NPRES.EQ.1) GO TO 400
          IF (NPRES.EQ.2) GO TO 400
          IF (NPRES.EQ.3) GO TO 400
          IF (NPRES.EQ.4) GO TO 400
          IF (NPRES.EQ.5) GO TO 400
          IF (NPRES.EQ.6) GO TO 400
          IF (NPRES.EQ.7) GO TO 400
          IF (NPRES.EQ.8) GO TO 400
          IF (NPRES.EQ.9) GO TO 400
          IF (NPRES.EQ.10) GO TO 400
C*****SUBJECT DID NOT FOLLOW DIRECTIONS
            CONTINUE
            WRITE (3,'(1X," ")')
          WRITE (3,'(1X,"   THE NUMBER YOU ENTERED IS NOT AVAILABLE")')
          GO TO 700
C
C
C*****GROUP 1A AND VIEWGRAPH NPRES ARE DETERMINED AND VALID
C*****DETERMINE FIRST IF SUBJECT IS FINISHED.  IF SO GO TO 500
400         CONTINUE
            IF (NPRES.EQ.10) GO TO 500
C
C*****NOT FINISHED WITH GROUP ANALYSIS. TITLE, DRAW BKGRND AND VGRAPH
```

```
      K=K+1
      CALL TITLE
      CALL BKGRND
      CALL VGRPH
C
C*****SUBJECT IS FINISHED WITH VIEWGRAPH, CLEAR SCREEN
C*****GIVE SUBJECT OPTION OF SEEING ANOTHER
C
      GO TO 10
C
C
C*****SUBJECT IS FINISHED ANALYZING THE GROUP
C*****SHOW SUBJECT THE NEXT GROUP
C
500   IX=IX+1
C*******DETERMINE IF TIMES NEED TO BE RECORDED
      IF (IX.NE.2) GO TO 18
      CALL ATIME
      CALL STIME
18    IF (IX.LT.4) GO TO 600
C*******WRITE OUT FINISH TIME AND NUMBER OF VIEWGRAPHS VIEWED.
      CALL WTIME
      K=K-1
      WRITE (6,39)K
39    FORMAT ('THE NUMBER OF GRAPHS VIEWED WAS   ',I2)
C*****SUBJECT HAS FINISHED ALL THREE GROUPS.
C*****RETURN TO PROGRAM DISSER TO END INTERACTIVE PART
      RETURN
C
C*****SUBJECT HAS ANOTHER GROUP TO ANALYZE.  OFFER IT.
600    CONTINUE
      WRITE (3,'(1X,"   YOU HAVE FINISHED THIS GROUPS ANALYSIS")')
      WRITE (3,'(1X,"   PLEASE GO ON TO THE NEXT GROUP")')
      RETURN
      END
      SUBROUTINE TITLE
C
C*****TITLE ALL VIEWGRAPHS
C
      COMMON/INPUT/IX, NPRES, K
       CALL JBEGIN
       CALL JDINIT (1)
       CALL JDEVON (1)
     CALL JDINIT (3)
     CALL JDEVON (3)
      IF (IX.GT.1) GO TO 50
       CALL JVSPAC (-1.,1.,-1.,1.)
       CALL JVPORT (-1.,1.,-1.,1.)
50     CALL JWINDO (0.,400.,0.,400.)
       CALL JOPEN
      CALL JSIZE (10.,10.)
      CALL JJUST (2,2)
      CALL JMOVE (200.,380.)
      CALL JCOLOR (3)
      CALL JFONT (3)
C
C*****MATCH TITLE AND VIEWGRAPH
100   IF(IX.NE.1) GO TO 300
      GO TO (110,120,130),NPRES
300   IF(IX.NE.2) GO TO 400
      GO TO (310,320,330),NPRES
400   IF(IX.NE.3) RETURN
      GO TO (310,310,310,320,320,320,330,330,330),NPRES
C
C*****SPECIFIC TITLE
110   CALL J3TEXT (31,31HSUMMER   DEMAND 33 CAMPUS BIKES)
      RETURN
120   CALL J3TEXT (32,32HSUMMER   DEMAND 33 3-SPEED BIKES)
      RETURN
130   CALL J3TEXT (31,31HSUMMER   DEMAND 33 DELUXE BIKES)
      RETURN
310   CALL J3TEXT (37,37HFALL FIGURES   DEMAND 67 CAMPUS BIKES)
      GO TO 315
320   CALL J3TEXT (38,38HFALL FIGURES   DEMAND 27 3-SPEED BIKES)
      GO TO 315
330   CALL J3TEXT (37,37HFALL FIGURES   DEMAND 25 DELUXE BIKES)
      GO TO 315
C*****ADD SUBTITLES TO FALL TITLES
```

```
 315    CALL JMOVE (230.,365.)
        CALL JSIZE (6.5,6.5)
        IF(IX.EQ.2) GO TO 340
 AUG 83 10:43:50    HARRIS FORTRAN 77 SAL OPTIMIZING COMPILER
ODULE NAME:  TITLE

        GO TO (344,345,346,347,348,349,350,351,352),NPRES
C*****FALL FIGURES WITHOUT SUMMER INVENTORY CONSIDERED
 340    CALL J3TEXT (31,31HSUMMER INVENTORY NOT CONSIDERED)
        RETURN
C*****FALL FIGURES WITH SUMMER INVENTORY CONSIDERED
 344    CALL J3TEXT (31,31HSUMMER PURCHASE 20 CAMPUS BIKES)
        RETURN
 345    CALL J3TEXT (31,31HSUMMER PURCHASE 40 CAMPUS BIKES)
        RETURN
 346    CALL J3TEXT (31,31HSUMMER PURCHASE NO CAMPUS BIKES)
        RETURN
 347    CALL J3TEXT (32,32HSUMMER PURCHASE 20 3-SPEED BIKES)
        RETURN
 348    CALL J3TEXT (32,32HSUMMER PURCHASE 40 3-SPEED BIKES.
        RETURN
 349    CALL J3TEXT (32,32HSUMMER PURCHASE NO 3-SPEED BIKES)
        RETURN
 350    CALL J3TEXT (31,31HSUMMER PURCHASE 20 DELUXE BIKES)
        RETURN
 351    CALL J3TEXT (31,31HSUMMER PURCHASE 40 DELUXE BIKES)
        RETURN
 352    CALL J3TEXT (31,31HSUMMER PURCHASE NO DELUXE BIKES)
        RETURN
        END
 C
        SUBROUTINE BKGRND
 C*****CREATE SHELL OF FIRST BARGRAPH
        DIMENSION X(2)
        CALL JPINTR (0)
        CALL JPEDGE (0)
        CALL JMOVE (40.,20.)
        CALL JRECT (40.,20.,200.,240.)
        CALL JMOVE (40.,100.)
        CALL JDRAW (200.,100.)
        CALL JMOVE (120.,20.)
        CALL JLSTYL (1)
        CALL JDRAW (120.,240.)
 C
 C*****CREATE SHELL OF SECOND BARGRAPH
        CALL JMOVE (230.,20.)
        CALL JRECT (230.,20.,380.,240.)
        CALL JMOVE (230.,100.)
        CALL JLSTYL (0)
        CALL JDRAW (330.,100.)
        CALL JMOVE (330.,20.)
        CALL JLSTYL (1)
        CALL JDRAW (330.,240.)
 C
 C*****CREATE TICKMARKS ND HORIZONTAL LINES ON SKELETON GRAPHS
        CALL JWIDE (15000)
        CALL JLSTYL (0)
        CALL JMOVE (41.,20.)
 C*****SET UP LOOP TO PUT TIK MARK ON FIRST BARGRAPH
        DO 20 I=1,21
      .    CALL JRMOVE (-2.,10.)
      .    CALL JRDRAW (2.,0.)
 20   .    CONTINUE
 C*****SET UP SAME LOOP TO PUT TIK MARK ON SECOND GRAPH
        CALL JMOVE (231.,20.)
        DO 30 I=1,21
      .    CALL JRMOVE (-2.,10.)
      .    CALL JRDRAW (2.,0.)
 30   .    CONTINUE
 C
```

```
C*****PUT IN HORIZONTAL DASHED LINES FOR E
C*****SET UP LOOP TO PUT IN DASHED LINES F
          CALL JLSTYL (1)
          CALL JMOVE (230.,20.)
          DO 40 I=1,3
      .      CALL JRMOVE (-100.,20.)
      .      CALL JRDRAW (100.,0.)
4 AUG 83 10:43:50    HARRIS FORTRAN 77 SAL OPTIM
MODULE NAME:  BAGRND

40    .      CONTINUE
C*****SET UP LOOP FOR UPPER DASHED LINES
          CALL JMOVE (230.,100.)
          DO 50 I=1,6
      .      CALL JRMOVE (-100.,20.)
      .      CALL JRDRAW (100.,0.)
50    .      CONTINUE
C*****DO THE SAME FOR THE SECOND BARGRAPH
          CALL JMOVE (380.,20.)
          DO 60 I=1,3
      .      CALL JRMOVE (-100.,20.)
      .      CALL JRDRAW (100.,0.)
60    .      CONTINUE
          CALL JMOVE (380.,100.)
          DO 70 I=1,6
      .      CALL JRMOVE (-100.,20.)
      .      CALL JRDRAW (100.,0.)
70    .      CONTINUE
C*****PUT IN BARGRAPH NOTATION
          CALL JSIZE (6.5,6.5)
          CALL JJUST (3,1)
          X(1)=37.
          X(2)=277.
          DO 80 J=1,2
      .      CALL JMOVE (X(J),20.)
      .      CALL JSTEXT (5,5H$1200)
      .      CALL JMOVE (X(J),40.)
      .      CALL JSTEXT (3,3H900)
      .      CALL JMOVE (X(J),60.)
      .      CALL JSTEXT (3,3H600)
      .      CALL JMOVE (X(J),80.)
      .      CALL JSTEXT (3,3H300)
      .      CALL JMOVE (X(J),100.)
      .      CALL JSTEXT (3,3H  0)
      .      CALL JMOVE (X(J),120.)
      .      CALL JSTEXT (3,3H300)
      .      CALL JMOVE (X(J),140.)
      .      CALL JSTEXT (3,3H600)
      .      CALL JMOVE (X(J),160.)
      .      CALL JSTEXT (3,3H900)
      .      CALL JMOVE (X(J),180.)
      .      CALL JSTEXT (4,4H1200)
      .      CALL JMOVE (X(J),200.)
      .      CALL JSTEXT (4,4H1500)
      .      CALL JMOVE (X(J),220.)
      .      CALL JSTEXT (4,4H1800)
      .      CALL JMOVE (X(J),240.)
      .      CALL JSTEXT (5,5H$2100)
80    .      CONTINUE
C*****LABEL GRAPHS

C*****BARGRAPH 1 LABEL
          CALL JSIZE (6.,6.)
          CALL JFONT (1)
          CALL JPATH (2)
          CALL JJUST (2,2)
          CALL JGAP (.2)
          CALL JMOVE (3.,170.)
          CALL JSTEXT (7,7HREVENUE)
          CALL JMOVE (0.,60.)
          CALL JSTEXT (5,5HCOSTS)
          CALL JMOVE (240.,170.)
          CALL JSTEXT (12,12HTOTAL PROFIT)
          CALL JMOVE (240.,60.)
          CALL JSTEXT (10,10HTOTAL LOSS)
          CALL JGAP (.0)
```

```
C
C*****SET UP GENERAL LEGEND
      CALL JFONT (1)
      CALL JPATH (1)
      CALL JJUST (1,1)
      CALL JSIZE (9.,9.)
C
C*****PUT IN LEGEND BOXES.  THIS MODEL HAS COLOR OR PATTERN
95    CALL JPINTR (1)
      CALL JPEDGE (0)
C******DRAW POLYGONS AND FILL THEM IN
      CALL JMOVE (10.,355.)
      CALL JPIDEX (2,22)
      CALL JRMOVE (0.,-15.)
      CALL JRRECT (10.,10.)
      CALL JPIDEX (1,23)
      CALL JRMOVE (0.,-15.)
      CALL JRRECT (10.,10.)
      CALL JPIDEX (4,24)
      CALL JRMOVE (0.,-15.)
      CALL JRRECT (10.,10.)
      CALL JPIDEX (3,25)
      CALL JRMOVE (0.,-15.)
      CALL JRRECT (10.,10.)
      CALL JPIDEX (7,26)
      CALL JRMOVE (0.,-15.)
      CALL JRRECT (10.,10.)
      CALL JPIDEX (5,27)
      CALL JRMOVE (0.,-15.)
      CALL JRRECT (10.,10.)
C
C*****PUT IN LEGEND TEXT THAT CORRESPONDS TO PATTERNED BOXES
      CALL JMOVE (33.,340.)
      CALL J3TEXT (33,33HTOTAL SALES PRICE    AT      /BIKE)
      CALL JMOVE (33.,325.)
4 AUG 83 10:43:50    HARRIS FORTRAN 77 SAU OPTIMIZING COMPILER
MODULE NAME:  BKGRND

      CALL J3TEXT (33,33HTOTAL PURCHASE COST  AT      /BIKE)
      CALL JMOVE (33.,310.)
      CALL J3TEXT (33,33HTOTAL SHORTAGE COST  AT      /BIKE)
      CALL JMOVE (33.,295.)


      CALL J3TEXT (33,33HTOTAL OVERSTOCK COST AT      /BIKE)
      CALL JMOVE (33.,280.)
      CALL J3TEXT (41,41HTOTAL WAREHOUSE COST AT      /BIKE OVER 40)
      CALL JMOVE (33.,265.)
      CALL J3TEXT (23,23HTOTAL CHANGE MODEL COST)
C
C*****BASIC SCREEN FOR ALL PRESENTATIONS
C*****RETURN TO SEE WHICH VIEWGRAPH GETS  PAINTED ON
      RETURN
      END
C
C
C
C
      SUBROUTINE VGRPH

C*****VIEWGRAPHS ASSOCIATED WITH EACH GROUP
C*****GO TO STATEMENT CONTROLS WHICH GROUP'S GRAPHS ARE DRAWN
C
C
      COMMON/INPUT/IX, NPRES, K
C******PUT IN SPECIFIC DATA IN TOP LEGEND BY CALLING LGND
      CALL LGND
       CALL JCCLOR (0)
      GO TO (100,300,400),IX
900   RETURN
C
C*****GROUP ONE DATA
100   CONTINUE
      GO TO (110,120,130),NPRES
C
```

```
C
C*****DATA FOR CAMPUS BIKES SUMMER DEMAND 33
C
 110      Y1=33.
          Y2=-13.
          Y3=-9.
          Y4=0.
          Y5=0.
          Y6=55.
          Y7=-27.
          Y8=-9.
          Y9=0.
          Y10=0.
          Y11=14.
          Y12=22.
          GO TO 500
C
C
C*****DATA FOR 3-SPEED BIKES SUMMER DEMAND 33
C
C
 120      CONTINUE
          Y1=67.
          Y2=-40.
          Y3=-13.
          Y4=0.
          Y5=0.
          Y6=110.
          Y7=-40.
          Y8=-5.
```



```
          Y9=-0.
          Y10=0.
          Y11=10.
          Y12=23.
          GO TO 500
C
C
C*****DATA FOR DELUXE MODEL SUMMER OPTION
C
C
 130      CONTINUE
          Y1=80.
          Y2=-40.
          Y3=-33.
          Y4=0.
           Y5=0.
          Y6=132.
          Y7=-30.
          Y8=-33.
          Y9=0.
          Y10=0.
          Y11=6.
          Y12=15.
          GO TO 500
C
C
C*****GROUP TWO DATA
C
C
 300      CONTINUE
          GO TO (310,320,330),NPRES
C
C*****CAMPUS BIKES FALL DATA ONLY
 310      CONTINUE
          Y1=100.
          Y2=-40.
          Y3=-4.
          Y4=-10.
          Y5=0.
          Y6=112.
          Y7=-34.
          Y8=-10.
          Y9=-10.
          Y10=0.
          Y11=47.
          Y12=32.
          GO TO 500
C
C*****3-SPEED BIKES FALL DATA ONLY
 320      CONTINUE
```

```
          Y1=87.
          Y2=-40.
          Y3=-10.
          Y4=C.
          Y5=C.
          Y6=90.
          Y7=-30.
          Y8=-9.
          Y9=C.
          Y10=0.
          Y11=13.
          Y12=1.
          GO TO 500
C
C*****DELUXE BIKES FALL DATA ONLY
330       CONTINUE
          Y1=80.
          Y2=-40.
          Y3=-14.
          Y4=C.
          Y5=C.
          Y6=100.
          Y7=-80.
          Y8=-80.
          Y9=C.
          Y10=0.
          Y11=27.
          Y12=-60.
          GO TO 500
C
C*****GROUP 3 DATA
C
C
400       CONTINUE
          GO TO (340,350,360,370,380,390,401,410,420),NPHES
C
C*****CAMPUS BIKES FALL DATA WITH 20 PURCHASE IN SUMMER
340       CONTINUE
          Y1=100.
          Y2=-40.
          Y3=-13.
          Y4=-10.
          Y5=C.
          Y6=134.
          Y7=-54.
          Y8=C.
          Y9=-17.
          Y10=C.
          Y11=41.
          Y12=64.
4 AUG 83 10:43:50   HARRIS FORTRAN 77 SAG OPTIMIZING COMPILER      0141
MODULE NAME:  VGNPM

          GO TO 500
C
C*****CAMPUS BIKE FALL (40/60) DATA WITH 40 PURCHASED IN SUMMER
350       CONTINUE
          Y1=78.
          Y2=-27.
          Y3=-9.
          Y4=-6.
          Y5=C.
          Y6=112.
          Y7=-40.
          Y8=-0.
          Y9=-12.
          Y10=0.
          Y11=30.
          Y12=60.
          GO TO 500
C
C*****CAMPUS BIKES FALL (60/50) DATA WITH NONE PURCHASED IN SUMMER
360       CONTINUE
          Y1=100.
          Y2=-40.
          Y3=-4.
          Y4=-10.
          Y5=-10.
          Y6=112.
```

```
      Y7=-53.
      Y8=-11.
      Y9=-17.
      Y10=-10.
      Y11=37.
      Y12=22.
      GO TO 500
C
C*****3-SPEED BIKES FALL (20,40) DATA WITH 20 PURCHASED IN SUMMER
370   CONTINUE
      Y1=67.
      Y2=-43.
      Y3=-27.
      Y4=0.
      Y5=0.
      Y6=134.
      Y7=-83.
      Y8=0.
      Y9=0.
      Y10=0.
      Y11=0.
      Y12=54.
      GO TO 500
C
4 AUG 83 10:43:50   HARRIS FORTRAN 77 SAU OPTIMIZING COMPILER     01414-00
MODULE NAME:  VGRPH

C*****3-SPEED BIKES FALL (20,40) DATA WITH 40 PURCHASED IN SUMMER
380   CONTINUE
      Y1=90.
      Y2=-40.
      Y3=0.
      Y4=0.
      Y5=0.
      Y6=90.
      Y7=-83.
      Y8=-14.
      Y9=-6.
      Y10=0.
      Y11=53.
      Y12=-10.
      GO TO 500
C
C*****3-SPEED BIKES FALL (20,40) DATA WITH NONE PURCHASED IN SUMMER
390   CONTINUE
      Y1=67.
      Y2=-43.
      Y3=-13.
      Y4=0.
      Y5=-13.
      Y6=90.
      Y7=-83.
      Y8=-9.
      Y9=0.
      Y10=-10.
      Y11=8.
      Y12=-9.
      GO TO 500
C
C*****DELUXE BIKES FALL (20,40) DATA WITH 20 PURCHASED IN SUMMER
401   CONTINUE
      Y1=80.
      Y2=-43.
      Y3=-48.
      Y4=0.
      Y5=0.
      Y6=140.
      Y7=-83.
      Y8=-11.
      Y9=0.
      Y10=0.
      Y11=-3.
      Y12=61.
      GO TO 500
C
C*****DELUXE BIKES FALL (0,20) DATA WITH 40 PURCHASED IN SUMMER
410   CONTINUE
```

```
            Y1=25.
            Y2=0.
            Y3=-43.
            Y4=0.
            Y5=0.
            Y6=100.
            Y7=-40.
            Y8=-11.
            Y9=0.
            Y10=0.
            Y11=-30.
            Y12=50.
            GO TO 500
C
C*****DELUXE BIKES FALL DATA WITH NONE PURCHASED IN SUMMER
420         CONTINUE
            Y1=20.
            Y2=-40.
            Y3=-14.
            Y4=0.
            Y5=-10.
            Y6=100.
            Y7=-80.
            Y8=-80.
            Y9=0.
            Y10=-10.
            Y11=12.
            Y12=-70.
            GO TO 500
C
C*****DRAW BARS WITH ASSIGNED DATA USING COLOR AND PATTERN
C
C
500         CONTINUE
            CALL JPINTR (1)
            CALL JPEDGE (0)
            CALL JMOVE (30.,100.)
            CALL JPIDEX (2,22)
            CALL JRRECT (10.,Y1)
            CALL JRMOVE (10.,0.)

            CALL JPIDEX (1,23)
            CALL JRRECT (10.,Y2)
            CALL JRMOVE (10.,0.)
            CALL JPIDEX (4,24)
            CALL JRRECT (10.,Y3)
            CALL JRMOVE (10.,0.)
            CALL JPIDEX (7,26)
            CALL JRRECT (10.,Y4)
            CALL JRMOVE (10.,0.)
            CALL JPIDEX (3,27)
4 AUG 83 10:43:50    HARRIS FORTRAN 77 SAU OPTIMIZING COMPILER
MODULE NAME:   OGRPH

            CALL JRRECT (10.,Y5)
C*****OVERSTOCK BARS
            CALL JMOVE (140.,100.)
            CALL JPIDEX (2,22)
            CALL JRRECT (10.,Y6)
            CALL JRMOVE (10.,0.)
            CALL JPIDEX (1,23)
            CALL JRRECT (10.,Y7)
            CALL JRMOVE (10.,0.)
            CALL JPIDEX (3,25)
            CALL JRRECT (10.,Y8)
            CALL JRMOVE (10.,0.)
            CALL JPIDEX (7,26)
            CALL JRRECT (10.,Y9)
            CALL JRMOVE (10.,0.)
            CALL JPIDEX (3,27)
            CALL JRRECT (10.,Y10)
C
C
C*****PUT IN PROFIT BARS
            CALL JMOVE (330.,100.)
             CALL JPIDEX (9,30)
            CALL JRRECT (10.,Y11)
            CALL JRMOVE (30.,0.)
             CALL JPIDEX (3,29)
            CALL JRRECT (10.,Y12)
C
```

```
C
C******PUT IN BOTTOM LABELS BY CALLING BTMLB SUBROUTINE
C
      CALL BTMLB
C
C
C*****ALLOW SUBJECT TO VIEW TOTAL VIEWGRAPH
      CALL JCLOSE
      CALL JPAUSE (1)
       CALL JEND
      RETURN
      END
C
C
C
C
      SUBROUTINE BTMLB
C*****PUT IN APPROPRIATE NUMBER PURCHASE LABELS IN EACH BAR GRAPH
C
C
       COMMON/INPUT/IX,NPRES, K
C*****SET UP STANDARD TEXT INFO.  THEN BRANCH TO PARTICULAR VG
C
      CALL JFONT (1)
      CALL JJUST (2,2)
      CALL JSIZE (6.,6.)
      CALL JCOLOR (3)
      CALL JMOVE (33.,230.)
C
C******BRANCH TO PUT IN 0/20, 20/40 OR 60/80 PURCHASE
      IF (IX.EQ.1) GO TO 100
      IF (IX.EQ.2.AND.NPRES.EQ.1) GO TO 300
      IF (IX.EQ.3.AND.NPRES.EQ.1) GO TO 300
      IF (IX.EQ.3.AND.NPRES.EQ.2) GO TO 350
      IF (IX.EQ.3.AND.NPRES.EQ.3) GO TO 350
      IF (NPRES.EQ.3) GO TO 375
      GO TO 100
C
C*****20/40 LABELS
100   CALL JSTEXT (6,6HBUY 20)
      CALL JMOVE (190.,230.)
      CALL JCOLOR (5)
      CALL JSTEXT (6,6HBUY 40)
      CALL JMOVE (335.,230.)
      CALL JCOLOR (6)
      CALL JSTEXT (6,6HBUY 20)
      CALL JMOVE (335.,230.)
      CALL JCOLOR (5)
      CALL JSTEXT (6,6HBUY 40)
      RETURN
C
C*****60/80 LABELS
300   CALL JSTEXT (6,6HBUY 60)
      CALL JMOVE (190.,230.)
      CALL JCOLOR (5)
      CALL JSTEXT (6,6HBUY 80)
      CALL JMOVE (335.,230.)
      CALL JCOLOR (6)
      CALL JSTEXT (6,6HBUY 60)
      CALL JMOVE (355.,230.)
      CALL JCOLOR (5)
      CALL JSTEXT (6,6HBUY 80)
      RETURN
C
C******40/60 LABELS
350   CALL JSTEXT (6,6HBUY 40)
      CALL JMOVE (190., 230.)
      CALL JCOLOR (5)
      CALL JSTEXT (6,6HBUY 60)
      CALL JMOVE (335.,230.)
      CALL JCOLOR (6)
      CALL JSTEXT (6,6HBUY 40)
      CALL JMOVE (355.,230.)
      CALL JCOLOR (5)
      CALL JSTEXT (6,6HBUY 60)
      RETURN
```

```
C
C******0/20 LABELS
375      CALL J3TEXT (9,8HBUY  0)
         CALL JMOVE (195.,230.)
         CALL JCOLOR (5)
         CALL J3TEXT (9,8HBUY 20)
         CALL JMOVE (335.,230.)
         CALL JCOLOR (5)
         CALL J3TEXT (9,8HBUY  0)
         CALL JMOVE (355.,230.)
         CALL JCOLOR (5)
         CALL J3TEXT (9,8HBUY 20)
         RETURN
C
         END
C
C
C
C
         SUBROUTINE LGND
C*****PUT IN APPROPRIATE REV, COST, DEMAND FIGURES IN LEGEND
C
         COMMON/INPUT/IX, NPRES, K
C
         CALL JFONT (1)
         CALL JPATH (1)
         CALL JJUST (1,1)
C*******WAREHOUSE COST
99       CALL JMOVE (245.,250.)
359       CALL JSIZE (9.,9.)
          CALL JCOLOR (7)
         CALL J3TEXT (4,4H$  5)
C*****CHANGE OVER COSTS
         CALL JSIZE (9.,9.)
         CALL JMOVE (245.,265.)
         CALL JCOLOR (3)
         CALL J3TEXT (4,4H$150)
         CALL JMOVE (245.,340.)
         CALL JCOLOR (2)
C*****CAMPUS BIKE FIGURES
         IF (NPRES.EQ.1) GO TO 100
         IF (IX.EQ.3.AND.NPRES.EQ.2) GO TO 100
         IF (IX.EQ.3.AND.NPRES.EQ.3) GO TC 100
C******3-SPEED BIKE FIGURES
         IF (NPRES.EQ.2) GO TO 200
         IF (NPRES.EQ.4) GO TO 300
         IF (NPRES.EQ.5) GO TU 200
         IF (NPRES.EQ.6) GO TJ 200
C******DELTAS AT DEFAULT
         GO TO 100
100      CALL J3TEXT (4,4H$ 25)
         CALL JMOVE (245.,325.)
         CALL JCOLOR (1)
         CALL J3TEXT (4,4H$ 10)
         CALL JMOVE (245.,310.)
         CALL JCOLOR (4)
         CALL J3TEXT (4,4H$  7)
         CALL JMOVE (245.,295.)
         CALL JCOLOR (3)
         CALL J3TEXT (4,4H$ 12)
299      RETURN
C
C
C*****3-SPEED BIKE FIGURES
200      CONTINUE
```

```
      CALL J3TEXT (4,4H$ 50)
      CALL JMOVE (245.,325.)
      CALL JCOLOR (1)
      CALL J3TEXT (4,4H$ 30)
      CALL JMOVE (245.,310.)
      CALL JCOLOR (4)
      CALL J3TEXT (4,4H$ 20)
      CALL JMOVE (245.,295.)
      CALL JCOLOR (3)
      CALL J3TEXT (4,4H$ 10)
      RETURN
C
C*****DELUXE BIKE FIGURES
C
300   CALL J3TEXT (4,4H$ 60)
      CALL JMOVE (245.,325.)
      CALL JCOLOR (1)
      CALL J3TEXT (4,4H$ 30)
      CALL JMOVE (245.,310.)
      CALL JCOLOR (4)
      CALL J3TEXT (4,4H$ 40)
      CALL JMOVE (245.,295.)
      CALL JCOLOR (3)
      CALL J3TEXT (4,4H$ 80)
      RETURN
      END
```

Appendix D

Color Treatment Identification

	CRT PENTEL Felt Pen ID	OUT PENTEL Felt Pen ID	RND PENTEL Felt Pen ID	PAS VENUS Pencil ID	B/W DI 3000 INTENSITY PATTERN ID and PATTERN
TOTAL SALES PRICE AT	104	104	134	16	22
TOTAL PURCHASE COST AT	119	119	109	18	23
TOTAL SHORTAGE COST AT	103	103	0	24	24
TOTAL OVERSTOCK COST AT	122	122	110	0	25
TOTAL WAREHOUSE COST AT	0	0	131	22	26
TOTAL CHANGE MODEL COST	30	30	107	17	27
UNDERSTOCK TOTAL	110	110	122	27	36
OVERSTOCK TOTAL	131	131	30	14	41

COLOR TREATMENT IDENTIFICATION

Appendix E

Consent Form

COLLEGE OF ENGINEERING AND APPLIED SCIENCES
Research Participation Agreement

Type of Research: PhD
Researcher: Mary S. McCully Phone: 897-7556
PLEASE READ THE FOLLOWING BEFORE YOU SIGN THE CONSENT FORM

Description of Procedure:

The purpose of this research is to gain a better working knowledge of a decision support system (DSS). A case study will be used to create a simulated decision making environment. The information needed is presented in a hardcopy pamphlet containing computer generated graphics sheets.

This study involves you taking the role of decision maker while interacting with computer generated graphics while completing part of the case study. No previous computer experience is needed. Initially you will complete a background experience questionnaire. Then a typed packet will be given to you that describes the case study scenario and contains the questions you need to answer in order to complete the experiment. You will solve each case study at your own pace. A post test will then be administered covering the information presented during the case studies. All questions are printed in hardcopy in a multiple choice format.

All other data collected during the experiment will be converted to statistical information with the understanding that confidentiality of such information will be maintained.

Thank you for your cooperation in this research.

CONSENT

My signature below, in return for the opportunity of participating as a subject in a scientific research investigation, hereby authorizes the performance upon me of the procedure described above. This consent I give voluntarily and after the nature and purpose of the experimental procedure, the known dangers, and the possible risks and complications have been fully explained to me. I knowingly assume the risks involved, and am aware that I may withdraw my consent and discontinue participation at any time without penalty to myself.

Signature:__________________________ Date:_______________

Effects of Alternative Chromatic Graphics Displays
in Decision Support Systems*

by

Mary S. McCully
U.S. Air Force

and

James E. Bailey
Arizona State University

* Materials in this paper were developed as part of Dr. McCully's Ph.D. dissertation.

Effects of Alternative Chromatics Graphic Displays in Decision Support Systems

ABSTRACT

This article describes a study that investigated the relative effectiveness on decision makers' performance when integrating color functionally into mixed text and graphics displays of Decision Support Systems. Five alternatives of color, 7 characteristics of decision makers, 3 levels of decision time and 6 levels of decision quality were the variables considered in the study. Sixty-three hypotheses were tested using the data collected from 120 subjects who had academic engineering backgrounds. The data contained insufficient evidence at the .05 significance level to conclude that color has a unique quality for functionally coding information in a DSS. If color is incorporated into a DSS, those colors currently preset on many CRTs appear best, but increased efficiency of decision makers can not be assumed.

KEYWORDS: computer graphics, MIS, color graphics, decision quality, decision time, information recall.

Effects of Alternative Chromatic Graphic Displays in Decision Support Systems

Introduction

Computer generated graphics supporting the option of color have become the heart of many decision support systems (DSS) in the industrial, business and military sectors. Technological advances, the intuitive appeal of color output and the increasing availability and usability of hardware/software have resulted in a steady growth in the application of color to information systems. However, these systems:

> have been growing seemingly in an ad hoc manner. While support and praise of these systems have been relatively high, it appears that efforts have fallen short of realizing maximum effectiveness and efficiency for the system designed (Dickson, Senn and Chervany, 1977, p. 913).

The technical development of the color graphical displays is ahead of reported, quantitative information regarding its contribution to the DSS. The challenge remains to ascertain its relative effectiveness on user performance.

Background

For over 25 years, color has been used as a method of information input coding. However, empirical evidence and guidelines on how to use it effectively are sparse and very narrow in scope. In fact, the literature has just begun to address how color impacts computerized decision support systems. The theoretical and the majority of the subjective literature are concurrent in their reports: color allows for relatively more efficient processing of information, thus improves performance over monochrome displays (De Mars, 1975; Friend, 1980; Miller, 1982; Morris, 1979; Truckenbroad, 1981).

The majority of the empirical studies based their conclusions on analysis of performance of the simple discrete tasks of search, locate, identify, and

count. The environments were typically isolated laboratory settings that often simulated an aircraft cockpit. The alternative codes to which color was compared for relative effectiveness included letters, numerals, geometric and abstract shapes. The subjects, whose performance constituted the data for analysis, were usually lacking in experience with the given task. Without exception the measure of performance was speed and/or accuracy. Findings in these experimental literature are not consistent with the theoretical or the majority of subjective reports. The empirical results can be summarized by stating the relative effectiveness of color is a function of the task undertaken. Generally, color was not found to have a unique quality for coding information in a display over other codes (Barker and Krebs, 1977; Christ, 1977; Christner and Ray, 1961; Cook, 1974; Tullis, 1981).

There continues to remain glaring gaps in the literature on the use of color in visual displays, especially those which do not involve search, count, locate and identify type of tasks. One of the fastest growing sectors of computer graphic users are managers and decision makers who are incorporating color graphics into their DSS (Berk, Brownston and Kaufman, 1982; McLamb, 1982). The question is, will color aid managers in their primary function of decision making. The subjective literature purports that color will speed up the decision making process significantly as well as improve recall or memory (Durrett, Zwiener and Freund, 1981; Friend, 1980; Miller, 1982; Morris, 1979). However, the question of how color affects speed and accuracy, two of the most coveted resources of decision makers, has not been thoroughly addressed. Dooley and Harkins (1970) investigated the effect color had on attention and memory, when incorporated into a graphic display that sat in the corner of the experiment room. The subjects were not aware the graphic display was a part of the experiment. The study presented insufficient evidence to conclude color

affected memory performance. Tullis (1981) had subjects perform a training and diagnostic task involving the status of telephone lines. The graphics were line schematics displayed on a cathode ray tube (CRT) interface. Conditions of the telephone lines were coded by either an achromatic shape or chromatic underlining bar (red, yellow, green). The task involved decision making but not memory. He concluded there was no significant differences in time or accuracy performance between the color underlining bar and the achromatic shapes.

There still remains much to be learned regarding when and how color can be used to increase a decision maker's performance. The purpose of this study was to investigate the relative effectiveness on decision maker's performance when integrating color text and graphic displays into typical decision support systems (DSS).

Conceptual Framework

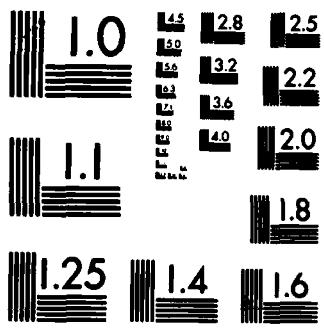
The DSS investigated in this study was composed of decision makers with academic engineering backgrounds. They solved a basic engineering case study which involved structured problems and deterministic data. The data were presented in mixed text and graphic displays as exemplified in Figure 1. Computer generated hardcopy was used to allow for a wide variety of colors. The decisional setting was one of managerial control of resources. The variables considered were classified as either treatment factors, characteristic factors or performance variables as delineated in Table 1.

The treatment factors included five alternative sets of colors. Each set was used to code the legend squares and column bars in each of the graphic displays. Traditional uses of color were utilized, that is green represented a revenue while red coded a cost. Colors of the same hue were not placed adjacent to each other. Color usage was compatable across all displays as a

AD-A145 561 EFFECTS OF ALTERNATIVE CHROMATIC MIXED GRAPHICS DISPLAYS IN DECISION SUPPORT SYSTEMS(U) AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH M S MCCULLY MAY 84 4/4

UNCLASSIFIED AFIT/CI/NR-84-46D F/G 9/2 NL

END
FILMED


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

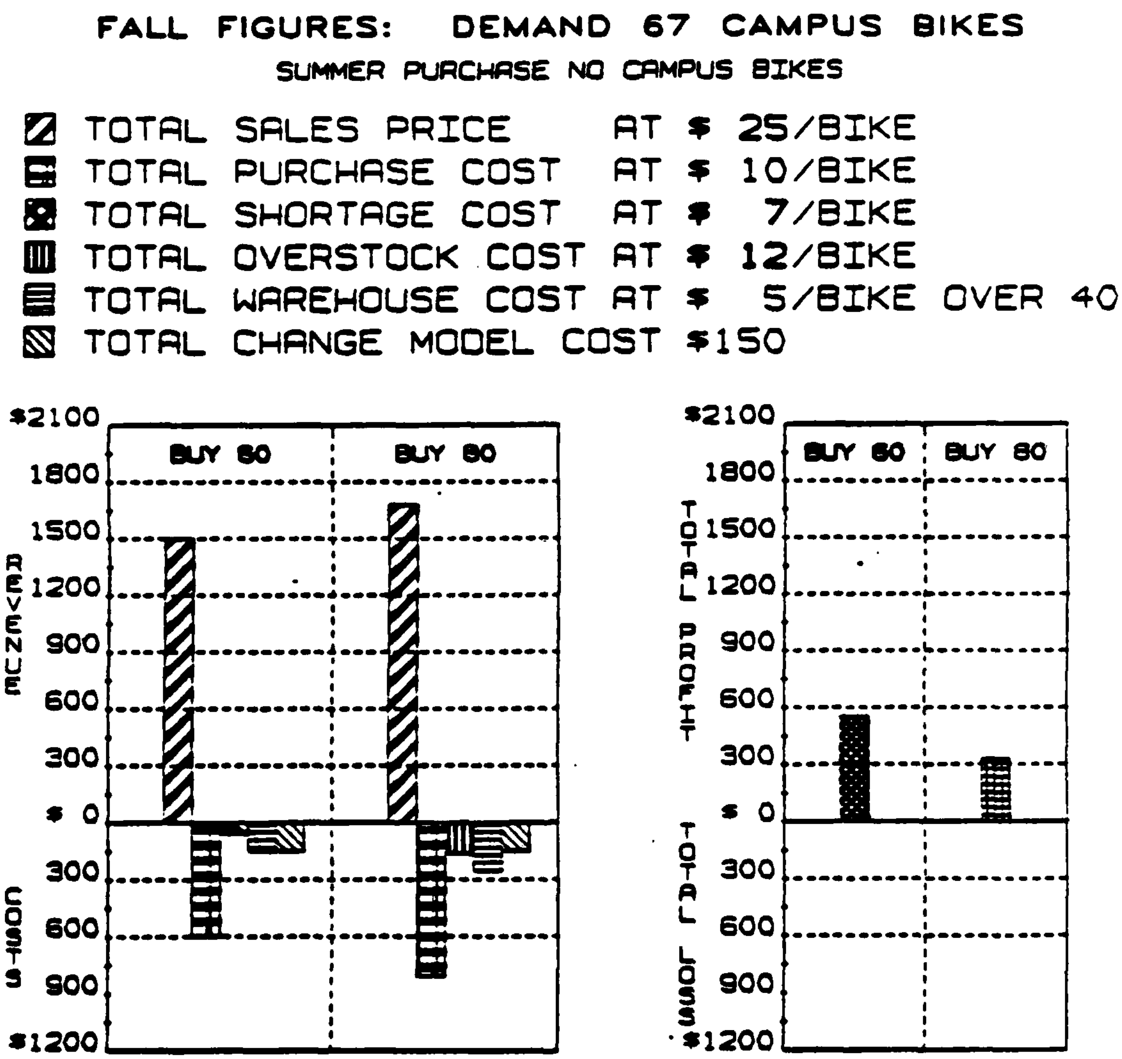


FIGURE 1 EXAMPLE OF MIXED GRAPHICS DISPLAY

Table 1

VARIABLES for ANALYSIS

Treatment Factor	Sample Size
Color	
CRT Color (CRT)	24
Black & White (B/W)	24
Outline (OUT)	24
Random (RND)	24
Pastel (PAS)	24

Characteristic Factor	Sample Size
Age (years)	
17-22	41
23-28	52
29+	27
Sex	
Female	33
Male	87
Years of College	
< 3	37
≥ 3 and < 5	30
≥ 5	53
Education Level	
Freshman/Sophomore	25
Junior/Senior	66
Graduate	29
CRT Experience	
Not a User (0-3 Never Use)	33
Moderate User (0-6 hrs/wk)	44
Frequent User (> 6 hrs/wk)	43
Attitude toward Computer Use	
Uncomfortable	31
Comfortable	48
Very Comfortable	41
Video Game Experience	
Do Not or Seldom Play (0-3 hrs/wk)	39
Moderate Player (> 3 hrs/wk)	81

Table 1 (cont.)

VARIABLES for ANALYSIS

Performance Variables	Sample Size
TIME (Seconds)	
Total Time	120
No Memory Time	120
Memory Time	120
QUALITY SCORE (%)	
Overall Score	120
No Memory Score	120
Combined Memory Score	120
Memory Only Score	120
Read Only Score	120
Extrapolate Information Score	120

particular color always symbolized the same information. The set of colors designated CRT were the primary colors preset on the majority of color CRTs. These included green, red, yellow, blue, turquoise, cyan, black, and white. The second set designated OUT was the same as the CRT except the legend squares and columns were only outlined in these colors, not solidly filled. This treatment was selected to see if the amount of color in the output had any affect. The third set, RND, consisted of a random selection of colors that are available, either primary or user specified, on the CRT colorwheel. This treatment was selected to investigate whether it was the typical stereotype colors CRT or just the presence of any color that had any affect on performance. This alternative included green, hot pink, white, turquoise, purple, orange, yellow and black. The fourth treatment, noted as PAS, used a random selection of pastel colors. The intensity of the color was significantly less than that in the CRT or RND alternative. The affect of color intensity was the test objective of this set which was made up of lime green, pink, yellow, white, purple, grey, blue and flesh. The monocrhome treatment, designated B/W, was used to investigate the most fundamental issue of whether the use of chromatic color in any of the tested configurations had an affect on decision making. For this alternative, eight of the patterns available on the DI-3000 graphics language package were used to generate the displays. The eight patterns selected were based on the criterion of non-similarity in pattern.

Seven characteristics of decision makers that may have influenced performance were included in the design. These were age, sex, years of college, education level, CRT experience, attitude toward the computer and experience interfacing with computers or video games. The specific levels and

sample size of each are shown in Table 1. These data were gathered from a questionnaire each subject completed prior to participating in the experiment.

The performance measured were time and quality of decision. These two variables were subdivided into 9 categories for more rigorous analysis of the research question. As indicated in table 1, the three levels of time represented the time taken to make: 1) those decisions which required no memory, 2) those decisions, some of which required memory of no longer available graphics and previous decisions and 3) all the decisions in the study. The decision quality variables measured performance ranging from the simplest decisions (direct reading, no computation and no memory) to the most difficult (computation , memory and extraction) as well as overall decision quality. Each level of decision quality is defined when the results are discussed. One hundred twenty observations were collected on each of the nine performance variables.

Method

The population of interest is decision makers who have an academic engineering background. Many of the subjects had been or were presently a member of the engineering workforce. An inventory control case study was selected as the content of the decision process. The problems were based on deterministic facts.

The case study consisted of four progressive scenarios. The background facts were given in text form. The decisions needed were made in conjunction with multiple choice questions. Information for only one scenario was available at a time, though previous facts and decisions were often needed to solve the currently worked scenario. The fourth scenario had no information

graphic displays. The average time spent completing the case study was approximately one hour.

The experiment's procedures and instrument were pretested using ten subjects. Based on the pretest, the procedure was modified to increase clarity. One hundred twenty subjects were selected and scheduled for the experiment. The treatments were completely randomized as to assignment to the subjects. The subjects were tested for color deficiency via the Dvorine Psuedo Isochromatic Test, and the experiment run.

Analysis

Hypotheses. Hypotheses were formulated that incorporated each of the performance variables with every combination of treatment and characteristic factor. Each null hypothesis stated that no difference existed between the mean performance of the five treatment factors listed in Table 1. Thus, rejection of the null hypothesis would indicate a significant difference between, at least, the worst and best performance, given the variability attributed to the specified characteristic factor had been taken into account. The generic null hypothesis

$$H(k)_0: \ \mu(i) = \mu(j) \qquad \text{for all } i,j \qquad \begin{array}{l} i = \text{CRT,RND,OUT,PAS,B/W} \\ j = \text{CRT,RND,OUT,PAS,B/W} \\ k = \text{Characteristic Factor} \end{array}$$

was tested for the 9 alternative measures of performance and 7 characteristic factors for a total of 63 different tests.

The analysis of the collected data was based on the appropriate statistics (F, Brown-Forsythe F, Welch W) derived from the 2-factor fixed effect ANOVA model. Correlation analysis of the time and decision quality performance variables resulted in insufficient evidence to conclude any form of these two

variables were correlated. The underlying criteria of independence, normality, homoscedasticity and additivity were rigorously examined to validate the appropriate model and statistic. The final results are consolidated in Table 2. Entries in this table are the p values of the models whose performance variable starts the row and whose characteristic factor head the column. Interpretation of the table would indicate, for example, that when the variability in performance due to age was accounted for one could conclude that a difference exists between the total time for at least the worst and best of the five color treatments with a significance level greater than or equal to .25.

For the models which contain more than one p value/entry, the criterion of additivity could not be assumed. These p values correspond to those calculated for each level of the characteristic factor under consideration and are reported in the same order as the levels are listed in Table 1. That is, for sex, the pairs contain the p values for female and male respectively.

The p value entries are highlighted for those models which the Duncan or Tukey-Kramer statistic indicated sufficient evidence to reject the null hypothesis of equal average performances across all five color treatments at the .05 significance level. A detailed discussion of these highlighted results follows.

Decision Time

Total Time represented the time (in seconds) needed to make those decisions where the basic problem facts and graphic displays were available for review. In eight of the nine models' the data contained insufficient evidence to conclude chromatic displays differentially affected the overall time required to make decisions. For those decision makers who had between three and five years college experience, decision time was shortest when the visual

Table 2

Analysis of Main Effect of Color

p Values

Reject Hypothesis if $p \leq .05$

Performance \ Factors	Color x Age	Color x Sex	Color x Col Yr	Color x Edu Lv	Color x CRT Ex	Color x Attit	Color x V Game
TIME (sec)							
Total	.25	.84	.79 .03* .50	.71	.55	.32	.70
No Memory	.86	.94	.75 .06 .81	.87	.86	.58	.94
Memory	.39	.24	.15	.39	.40	.12	.44
QUALITY SCORE (%)							
Overall	.19	.01* .12	.22	.26	.15	.13	.03*
No Memory	.24	.13	.13	.11 .06 .61	.43	.21	.06
Combined	.32	.04* .17	.36	.55	.32	.36	.11
Memory Only	.19	.06 .20	.15	.04*	.38	.10	.05*
Read Only	.53	.20 .06	.38	.92	.26	.25	.22
Extrapolate	.64	.28	.23	.18	.11	.21	.20

*Rejected hypothesis of $p \leq .05$

display was functionally colored with the RND alternative. The B/W, PAS, CRT and OUT were the remaining rank ordering as shown in Table 3. RND outscored B/W by 2.7 percent, while B/W shortened total time by 11.5 percent over CRT. RND and B/W were not significantly different from each other, nor were B/W, PAS, CRT and OUT.

No memory time was the elapsed seconds when formulating decisions which required no memory but extraction of information directly available in the output displays. Results from analyses were consistent across every case: the use or lack of color did not affect no memory decision time across the five treatments when compared at the .05 level of significance.

Memory time was the seconds spent when making decisions which sometimes required memory of past displays and sometimes required no memory. Analysis overwhelmingly concluded the use of a color did not significantly affect memory time.

Decision Quality Score

Overall Score results gave sufficient evidence to conclude that color influenced the decision quality in only two cases. The overall decision quality of females, as shown in Table 4, were ranked best to worst as CRT, B/W, PAS, RND and OUT. Analysis showed insignificant differences between the first four and between the last four treatments. The CRT treatment yielded 16.5 percent increase in overall score over the B/W treatment and a 46 percent increase over the OUT treatment. These results did not hold for males. When video game experience was accounted for, the ranking of the overall scores was CRT, B/W, RND, OUT, and PAS as shown in Table 5. The first four were not significantly different from one another, nor were the last three. The CRT treatment yielded only a .7 percent increase in overall score over B/W and a 12 percent increase over PAS.

Table 3

ANALYSIS of COLOR for 3 to 5 YEARS COLLEGE EXPERIENCE vs TOTAL TIME

Performance Not Equal

Across All Treatments at $p \leq .05$

Color Factor / Performance	RND BEST	B/W	PAS	CRT	OUT WORST
Total Time (Secs)	1826.6	1878.0	1934.0	2093.5	2339.5
TUKEY-KRAMER GROUPINGS	A	A			
		B	B	B	B

Table 4

ANALYSIS OF COLOR for FEMALES vs OVERALL DECISION SCORE

Performance Not Equal

Across All Treatments at $p \leq .05$

Color Factor / Performance	CRT BEST	B/W	PAS	RND	OUT WORST
Overall Score (%)	77.6	66.6	66.1	61.2	53.3
TUKEY-KRAMER GROUPINGS	A	A	A	A	
		B	B	B	B

Table 5

ANALYSIS of COLOR x V GAME EXPERIENCE vs OVERALL DECISION SCORE

Performance Not Equal

Across All Treatments at $p \leq .05$

Color Factor	CRT	B/W	RND	OUT	PAS
Performance	BEST				WORST
Overall Decision (%)	68.9	68.4	67.8	65.3	61.6
DUNCAN MULTIPLE RANGE GROUPINGS	A	A	A	A	
			B	B	B

Table 6

ANALYSIS of COLOR x FEMALE vs COMBINED SCORE

Performance Not Equal

Across All Treatments at $p \leq .05$

Color Factor	CRT	B/W	PAS	RND	OUT
Performance	BEST				WORST
Combined Score (%)	78.8	68.8	67.5	64.4	53.3
TUKEY-KRAMER GROUPINGS	A	A	A	A	
		B	B	B	B

In general, overall decision quality was not affected by the existence of color. When there was a significant difference, the rankings were similar with chromatic CRT the best but not significantly better than monochrome B/W, the second best. The actual average increase of CRT score over B/W was 8.6 percent.

No Memory Scores represented decisions in which all information was available at the time the decisions were made. The analysis indicated insufficient evidence to conclude the color treatment had a significant impact on any of the tested circumstances.

Combined Memory Score represented decisions requiring information, some of which was available at the time of the decision as well as memory of previously presented information. All models except female performance showed indifference to the color treatment. Females performed these combined memory decisions best to worst when interfacing with CRT, B/W, PAS, RND, and OUT respectively as shown in Table 6. Analysis indicated insignificant groupings between the first four as well as the last four treatments. However, females performed significantly better with the CRT solid filled colors than with the same colors in outline OUT form. Females using a chromatic CRT display outscored those females interfacing with a monochrome B/W display by 14.5 percent for the combined memory decisions. This difference was not statistically significant.

Memory Only Score evaluated decisions after all graphic displays were shown but were no longer available for reference. These decisions had to be made based solely on the memory of the decision maker. The analysis indicated only two of the eight models contained evidence to suggest the color of the graphic display affected memory only decision performance: those which accounted for the variability due to education level and for video game

experience. As indicated in Table 7 and 8, the ranking of treatments was in each: CRT, B/W, OUT, PAS and RND. The first three treatments formed an insignificantly different group as did the last four. After the variability in performance attributed to education level or video game experience was accounted for, chromatic CRT outscored monochrome B/W by approximately 4 percent and 7.8 percent respectively. This represented an average increase of approximately 5.9 percent in memory only score using CRT over B/W. However, these differences were not statistically significant at the .05 level. The random RND chromatic treatment ranked last, with an average of 17 percent decrease in memory only score.

Read Only Score represented decisions that required only reading information to get the decision directly from the graphic displays. Neither memory or extrapolation of the information was necessary. None of the eight models investigated suggested color differentially affected how the decision maker processed the graphical information.

Extrapolation Score was composed of the decisions which required manipulation or forecasting of the information presented in one or more graphic displays. Analyses overwhelmingly indicated insufficient evidence to conclude the colors, hues or amount of color differentially affected how the decision maker reached these types of decisions.

Summary

Of the 25 models that were examined whose performance was a function of decision time, all but one resulted in insufficient evidence to conclude that the specific colors, the hues or the amount of color in a mixed graphic display significantly affected decision time at the .05 level. The single case for which analyses indicated a difference in decision time attributed to color is shown in Table 9. The best chromatic display RND decreased decision time by

Table 7

ANALYSIS of COLOR x EDUCATION LEVEL vs MEMORY SCORE

Performance Not Equal

Across All Treatments at $p \leq .05$

Color Factor / Performance	CRT BEST	B/W	OUT	PAS	RND WORST
Memory Score (%)	67.4	62.6	61.3	59.6	57.7
TUKEY-KRAMER GROUPING	A	A	A		
		B	B	B	B

Table 8

ANALYSIS of COLOR x V GAME EXPERIENCE vs MEMORY SCORE

Performance Not Equal

Across All Treatments at $p \leq .05$

Color Factor / Performance	CRT BEST	B/W	OUT	PAS	RND WORST
Memory Score (%)	67.5	62.6	61.3	59.6	57.7
DUNCAN MULTIPLE RANGE GROUPINGS	A	A	A		
		B	B	B	B

Table 9

RANKING of TREATMENTS

by

PERFORMANCE

Groupings of Duncan/Tukey-Kramer

$\alpha = .05$

Color Ranking Characteristic vs Performance	BEST 1	2	3	4	WORST 5
TIME					
Col Ex vs Total Time	RND	B/W	PAS	CRT	OUT
DECISION QUALITY					
Female vs Overall Score	CRT	B/W	PAS	RND	OUT
V Game vs Overall Score	CRT	B/W	PAS	OUT	RND
Female vs Combined Mem	CRT	B/W	PAS	RND	OUT
Edu Lv vs Memory Only	CRT	B/W	OUT	PAS	RND
V Game vs Memory Only	CRT	B/W	OUT	PAS	RND

only 2.7 percent over the monochrome B/W while B/W showed a decrease in decision time by 11.5 percent over the industry standard chromatic scheme of CRT. These results are inconsistent with the theoretical and the majority of the subjective literature which contend color significantly speed the information processing, thus decision time.

Of the 42 models that were examined whose performance was a function of quality of decision, the majority 37 of them resulted in insufficient evidence to conclude the specific colors, hue or amount of color of the mixed graphic display significantly affected performance at the .05 significance level. These results are inconsistent with the theoretical and subject literature, which purport color when integrated into information display allows for better decisions. However, these findings are concurrent with the objective literature which was summarized by stating chromatic color does not appear to have a unique quality for coding *information* over other achromatic codes (Christ and Corso, 1983). For some tasks it may be beneficial, in others it may be detrimental. In this research, chromatic color did not influence performance in a way different than the achromatic code for a majority of the models. Where it did, the influence was always in the positive direction as it consistently ranked CRT better than achromatic B/W, though never with an interval which was statistically significant at the .05 level.

Unlike any other study reported in the literature, this research employed several chromatic alternatives in the experimental design. This was done to investigate if it was just the presence of color that affected performance or a particular set. As discussed previously, decision quality performance was not affected differently by any of the five alternatives. Based on these results, it can be concluded that despite the color set used, the hue or the amount of

color, performance was not affected differently than when a monochrome intensity pattern was used.

In the five models whose data contained sufficient evidence which suggested chromatic color affected decision quality performance, a pattern was detected as displayed in Table 9. In all five models, the chromatic CRT alternative was best, indicating an average of approximately 8.7 percent increase in decision quality score over the consistently second place monochrome B/W display. This difference was never statistically significant. Based on these results, it is concluded that if a color alternative is chosen those colors preset on the majority of CRTs are the ones which result in the best decision quality. However, since new color alternatives are becoming available for graphics displays and not all color alternatives were investigated, the generalizaton of this conclusion is limited.

Overall Conclusions Generalized to Problem Statement

The primary objective of this research was to investigate the relative effectiveness on performance when functionally integrating color into a DSS whose decision maker has an academic engineering background and solves structured problems based on deterministic data presented in a mixed graphics form on a computer generated hardcopy in a managerial role of control of resources.

Based on the results of 57 of 63 tested hypotheses, the attribute of color when functionally integrated in a mixed graphics hardcopy display did not significantly affect performance. In these test cases, quality of decision was not improved or degraded significantly at the .05 significance level, nor was decision time.

In the 6 models where the data contained evidence that the color alternative affected performance, the CRT alternative ranked best in 5

instances, all which were a function of decision quality. The average improvement was approximately 8.7 percent over the monochrome. Specifically, females performed best when interacing with the chromatic CRT alternative, but it was never a significant improvement over the monochrome B/W intensity pattern. In the other model where a difference in performance was suggested, the chromatic RND displays ranked first, with the monochrome B/W ranking second relative to decision making time. The improvement was significant with respect to the monochrome. Monochrome intensity pattern showed a 11.5 percent improvement in performance time over the CRT alternative.

In general, these results are concurrent with other empirical studies on the effects of color when integrated into visual displays. The data do not contain sufficient evidence to conclude the chromatic colors have a unique quality for functionally coding information in DSS over a monochrome intensity pattern. However, if color is functionally added to a hardcopy mixed graphics display, the solid filled intense hue of the preset colors available on a majority of the current CRTs appear to be the best overall. The manager should not assume increased efficiency in a decision maker's performance because of the functional color attribute.

<u>BIBLIOGRAPHY</u>

Barker, E. & Krebs, M. Color effects on human performance: An annotated bibliography. Technical Report ONR-CR213-136-1f. Minneapolis, Minnesota: Honeywell, Inc., April 1977.

Belie, R. and Rapagnani, N. Color Computer Graphics - a design and analysis potential scarcely tapped. Astronautics and Aeronautics, June 1981, 19, 41-48.

Benbasat, I. and Taylor, R. The impact of cognitive sytles on information system design. Management Information Systems Quarterly, June 1978, 2, 43-53.

Benbasat, J. and Schroeder, R. An experimental investigation of some MIS design variables. Management Information Systems Quarterly, March 1977, 1, 37-49.

Berenson, M., Levine, D. and Goldstein, M. Intermediate statistical methods and applications: A computer package approach. New Jersey: Prentice Hall, Inc., 1983.

Berk, T., Brownston, L. and Kaufman, A. A new color-naming system for graphics languages. IEEE Computer Graphics and Applications, May 1982, 2, 37-42.

Box, G., Hunter, W. and Hunter J. Statistics for Experimentation. New York: John Wiley and Sons, Inc., 1978.

Brown, M. and Forsythe, A. The small sample behavior of some statistics which test the equality of several means. Technometrics, February 1974, 16, 129-132.

Christ, R. Review and analysis of color coding research for visual displays. Human Factors, 1975, 17, 542-570.

Christ, R. Four years of color research for visual displays. Proceedings of The Human Factors Society, 21st annual meeting. San Francisco, California, October, 1977, 319-321.

Christ, R. and Corso, G. The effects of extended practice on the evaluation of visual display codes. Human Factors, February 1983, 25, 71-84.

Christner, C. and Ray, H. An evaluation of selected combinations of target and background coding on map-reading performance - experiment V. Human Factors, 1961, 3, 131-146.

Cook, T. Colour coding: A review of the literature. NTIS No. AD A001555, November 1974.

CRT graphics terminal market to exceed $1.4 billion by 1986. Interface Age, February 1982, 7, 19.

Demars, S. Human factors considerations for the use of color in display systems. NASA-TM-X-72196, TR-1329, January 1975. NTIS N75-14446.

DI-3000 user manual. Colorado: Precision Visuals, Inc., 1979.

Dickson, G., Senn, J. and Chervany, N. Research in management information systems. Management Science; May 1977, 23, 913-923.

Dooley, R. and Harkins, L. Functional and attention-getting effects of color on graphics communications. Perceptual and Motor Skills, 1970, 31, 851-854.

Durrett, H. Color display systems: The state of the art. Behavior Research Methods and Instrumentation, 1979, 11, 127-130.

Durrett, J. and Trezona, J. How to use color displays effectively, April 1982, BYTE 7, 50-53.

Durrett, J. and Zwiener, C. A microcomputer-based laboratory: Real world experience. Behavior Research Methods and Instrumentation, April 1981, 13, 209-212.

Durrett, J., Zwiener, C. and Freund, R. The human factors of color display-based instruction. Proceedings of the National Educational Computing Conference, 1981, 284-286, (ERIC 207 526).

Dvorine, I. Color perception testing charts (second edition). Baltimore, Maryland: Scientific Publishing Company, 1944.

Friend, D. Color graphics information systems boost productivity. Mini-Micro Systems, May 1980, 13, 181-182+.

Grandjean, E., Hunting, W. and Pidermann, M. VDT workstation design: Preferred sittings and their effects. Human Factors, April 1983, 25, 161-175.

Haber, R. and Wilkinson, L. Perceptual components of computer displays. IEEE Computer Graphics and Applications, May 1982, 2, 23-26.

Hanson, D. Introduction to the special issue on the application of color to displays. Proceedings of the Society of Information Display, first quarter 1979, 20, 1-2.

Keister, R. Data entry performance on color versus monochromatic displays. Proceedings, Human Factors Society. Rochester, New York, October 1981, 736-740.

Kelley, N. Computer graphics: Info at a glance. Infosystems, December 1979, 26, 37-39.

Kendall, M. and Stuart, A. The advanced theory of statistics. London: Griffin and Company Limited, 1958.

Leith, L. and O'Shea, F. Microcomputer graphics in education. European Association for Microprocessing and Microprogramming Journal, July 1980, 6, 205-208.

Levene, H. Robust tests for equality of variances in contributions to probability and statistics. Olkin, I. (editor). California: Stanford University Press, 1960.

Lucas, H. Performance and the use of an information system. Management science, April 1975, 21, 908-919.

Lucas, H. An experimental investigation of the use of computer-based graphics in decision making. Management Science, July 1981, 27, 757-768.

Lucas H. and Nielson, N. The impact of information presentation on learning and performance. Management Science, October 1980, 26, 982-993.

Mason, R. and Mitroff, J. A program for research in management information systems. Management Science, January 1973, 19, 475-485.

McCully, M. Effects of alternative chromatic mixed graphics displays in decision support systems. Dissertation. Arizona State University, May 1984.

McLamb, K. Persuasion made easy with powerful presentations. Personal Computing, February 1982, 6, 35-39.

Miastkowski, S. and Wrege, R. Computer graphics. Popular Computing, December 1981, 1, 70-75.

Miller, F. CRT terminals get smarter, cheaper. Infosystems, September 1981, 28, 100-102.

Miller, F. The world of CRT terminals. Infosystems, June 1982, 29, 86-92.

Miller, I. Computer graphics for decision making. Harvard Business Review, November-December 1969, 47, 121-132.

Morris, J. Using color in industrial control graphs. Control Engineering, July 1979, 26, 41-45.

Neter, J. and Wasserman, W. Applied linear statistical models. Illinois: Richard Irwin, Inc., 1974.

Ramsey, H., Atwood, M. and Kirschbaum, P. A critically annotated bibliography of the literature on human factors in computer systems. Technical Report SAI-78-070-DEN. Englewood, Colorado: Science Applications, Inc., May 1978.

Riesenfeld, R. Current trends in computer graphics. Computers and Graphics, 1978, 3, 115-122.

Roark, D. Toward realtime interactive color graphics. Computer Design, July 1982, 21, 167-173.

Takeuchi, H. and Schmidt, A. New promise of computer graphics. Harvard Business Review, January-February 1980, 58, 122-131.

Teichner, W. Color and visual information-coding. Proceedings of the Society for Information Display, 1979, 20, 3-9.

Teichner, W., Christ, R. and Corso, G. Color research for visual displays. Technical Report ONR-CR213-102-4F. Las Cruces, New Mexico: New Mexico State University, Department of Psychology, June 1977.

Teichner, W. and Krebs, M. Visual search for simple targets. Psychological Bulletin, 1974, 81, 15-28.

Truckenbroad, J. Effective use of color in computer graphics. Computer Graphics (ACM), August 1981, 15, 83-90.

Tullis, T. An evaluation of alphanumeric, graphic and color information displays. Human Factors, October 1981, 23, 541-550.

Wallace, B. Siggraph. Creative Computing, January 1981, 7, 44-49.

Waller, R., LeFrere, P. and McDonald-Ross, M. Do you need that second color? IEEE Transactions on Professional Communication, June 1982, 25, 80-85.

Zmud, R. Individual differences and MIS success: A review of the Empirical literature. Management Science, October 1979, 25, 966-979.

END

FILMED

10-84

DTIC